



DEME
ANDREA

BESZÉD • KUTATÁS • ALKALMAZÁS

MAGÁNHANGZÓK
EJTÉSE ÉS ÉSZLELÉSE
A SZOPRÁN-
ÉNEKLESBEN

 ELTE
EÖTVÖS
KIADÓ

Deme Andrea

MAGÁNHANGZÓK EJTÉSE ÉS ÉSZLELÉSE
A SZOPRÁNÉNEKLÉSBEN

Beszéd • Kutatás • Alkalmazás

Deme Andrea

MAGÁNHANGZÓK
EJTÉSE ÉS ÉSZLELÉSE
A SZOPRÁNÉNEKLÉSBEN

Budapest, 2016



A kötet megjelenését a Magyar Tudományos Akadémia támogatta.



Lektorálták:
Bárkányi Zsuzsanna
Bóna Judit
Vicsi Klára

© Deme Andrea, 2016

ISBN 978-963-312-261-7

ISSN 2064-4442

A borítón: Maria Callas görög-amerikai szoprán operaénekesnő (1923–1977) 1958-ban, Violetta szerepében (La Traviata). Fotó: Houston Rogers.



**E Ö T V Ö S
K I A D Ó**

www.eotvoskiado.hu



Felelős kiadó: Hunyady András, ügyvezető igazgató

Felelős szerkesztő: Gaborják Ádám

Projektvezető: Sándor Júlia

Tipográfia: Bornemissza Ádám

Borító: Csele Kmotrik Ildikó

Nyomdai kivitelezés: Multiszolg Bt.

Tartalom

Sorozatszerkesztői előszó	9
Előszó	11
1. Általános bevezetés	15
1.1. A magánhangzók produkciója és percepciója	15
1.1.1. A beszédkeltés folyamata, különös tekintettel a magánhangzók produkciójára ...	15
1.1.1.1. A zöngé (azaz a forrás) keletkezése, az alaphfrekvencia	16
1.1.1.2. A toldalékcso mint akusztikus szűrő, a formánsok	21
1.1.1.3. A magánhangzók artikulációs és akusztikai jellemzése	28
1.1.1.4. A magyar köznyelvi magánhangzók artikulációs és akusztikai jellemzése ...	33
1.1.2. A magánhangzók észlelése	38
1.1.2.1. Beszédpercepciós elméletek	39
1.1.2.2. A magánhangzó-azonosítás kísérletes eredményei és a magánhangzószerű hangok feldolgozása a percepciós rendszerben ...	42
1.1.2.3. A magyar magánhangzók azonosításának kísérletes eredményei	48
1.2. Az éneklés és az énekelt magánhangzók fonetikai jellemzése	51
1.2.1. Az énekelt magánhangzók artikulációs vizsgálata	55
1.2.1.1. A vertikális és horizontális gégehelyzet énekléskor	55
1.2.1.2. A gége belső mozgásai: a nyelvcsont szerepe az éneklésben.	60
1.2.2. Az énekelt magánhangzók akusztikai vizsgálata: rezonanciák a szopránéneklésben	62
1.2.3. Az énekelt magánhangzók észlelése	72
1.2.3.1. Az énekelt magánhangzók észlelésével kapcsolatos gyakori feltételezések	73
1.2.3.2. Az énekelt magánhangzók észlelésének vizsgálata	74
1.3. A jelen kutatás célja, kérdések, hipotézisek	92
2. A magyar magánhangzók ejtése és észlelése a magas alaphfrekvenciájú szopránéneklésben	95
2.1. Anyag, módszer és kísérleti személyek	95
2.2. Eredmények	97
2.2.1. Észlelési vizsgálat	97

2.2.2. Akusztikai elemzés	102
2.3. Következtetések	106
3. A mássalhangzó-környezet hatása az énekelt magánhangzók észlelésére: a mássalhangzó képzésmódjának hatása	111
3.1. Anyag, módszer és kísérleti személyek	113
3.2. Eredmények	114
3.3. Következtetések	118
4. A mássalhangzó-környezet hatása az énekelt magánhangzók észlelésére: azonosítás mássalhangzó-környezetben, izolált ejtésben, valamint a zöngéindítás mint akusztikai kulcs hiánya esetén	123
4.1. Anyag, módszer és kísérleti személyek	127
4.2. Eredmények	129
4.3. Következtetések	133
5. A mássalhangzó-környezet hatása az énekelt magánhangzók észlelésére: a magánhangzók azonosítása értelmes és értelmetlen szavakban énekes és nem énekes hallgatók percepciójában	139
5.1. Anyag, módszer és kísérleti személyek	142
5.2. Eredmények	145
5.2.1. Bevezető gondolatok a résztvevők benyomásai alapján	145
5.2.2. Az ejtési szándéknak megfelelő magánhangzó-azonosítás az értelmes és az értelmetlen szavakban	146
5.2.3. A magánhangzók ejtési szándéknak megfelelő azonosítása a két kondícióban	147
5.2.4. A tévesztések tendenciái az egyes magánhangzók esetében az alaphérvencia függvényében	149
5.3. Következtetések	158
6. Az énekelt magánhangzók akusztikai szerkezete néhány módszertani újítás tükrében . . .	165
6.1. Az elektrogloottográfia és az elektrogloottogram	166
6.2. Az inverz szűrés, az áramlási gloottogram, valamint az elektrogloottográfia szerepe az inverz szűrésben	169
6.3. A kísérlet anyaga és módszere, valamint a kísérleti személyek	170
6.3.1. Felvét elkészítés, az akusztikai vizsgálat anyaga	170
6.3.2. Akusztikai elemzések	174
6.4. Eredmények	176
6.4.1. Az énekelt magánhangzók első formánsának alakulása	177
6.4.2. Az énekelt magánhangzók második formánsának alakulása	178

6.4.3. Az akusztikai magánhangzótér az alapprofrekvencia függvényében	179
6.5. Következtetések	180
7. A szopránénekülésben megjelenő magánhangzók észlelése az akusztikai szerkezet tükrében	185
7.1. Traunmüller modellje és a modell lehetséges predikciói az énekelt magánhangzók azonosítására	186
7.2. Anyag, módszer, kísérleti szenélyek	190
7.3. Eredmények	191
7.4. Következtetések	202
8. Összegzés és következtetések	209
8.1. A hipotézisek vizsgálata	209
8.2. A kötetben bemutatott kutatómunka fontosabb eredményeinek rövid összefoglalása	215
8.3. A vizsgálatok közvetlen és közvetett hozadékai, kitekintés	217
9. Irodalom	221
Production and perception of Hungarian vowels in high-pitched soprano singing	231

Sorozatszerkesztői előszó

A (magas alaphfrekvenciájú) szoprán éneklésben a magánhangzók azonosítása a hallgatók számára nem könnyű feladat. Ennek az az oka, hogy a magánhangzó artikulációjának bizonyos értelemben ellentmond a magas alaphang létrehozására való törekvés az énekes részéről (ennek részleteiről és magyarázatáról lásd a jelen kötet megfelelő fejezeteit).

A magyar köznyelvi magánhangzóknek a szoprán éneklésben jellemző sajátosságos megvalósulásának leírására e könyv szerzőjének kutatásai előtt még nem született tudományos igényű vizsgálat. Ugyanakkor az adott nyelv hangzókészletét is figyelembe vevő, jól kontrollált kísérletei révén ez a könyv az énekelve ejtett magánhangzók észlelésének nemzetközi szakirodalmában is újdonságot hoz.

A szerző alaposan utánajárt a szakirodalomban olvasható megállapítások ellentmondásainak, nagy szorgalommal és alaposággal törekedett az olykor következtetlenül kidolgozott kísérleti módszertanokból előállt fél igazságok helyreigazítására. Saját kísérleteit a korábbi kutatások tanulságaiból okulva, ugyanakkor módszertani innovációkat is bevezetve végezte el. Az itt bemutatott kutatássorozat mindezek miatt nemcsak világszínvonalú, de a témakörben a létező legkorszerűbb tudást mutatja be az olvasónak.

Az elméleti háttér ismertetése, a kísérletek egymásra következő bemutatása logikus és fokozatos. A szerző az artikuláció, az akusztikum és a percepció hármasságában bontja ki a befogadáshoz szükséges információkat. Az egyik tudományos kérdésre kapott válasz alapján megfogalmazódó következő kérdésre keres választ a soron következő kísérlet, így haladunk előre a megismerés folyamatában.

Azt gondolhatnánk, hogy az itt tárgyalt kérdéskör specifikus volta miatt igen keveseket érdekelhet. E könyv azonban nagyon sok alapvető tudományos eredményt foglal össze, amelyeket korábban magyarul nemigen olvashattunk. Így nemcsak a szaktudomány művelőinek, de az érdeklődőknek és a tanulni vágyóknak is olyan mennyiségű és minőségű ismeretet ad át a beszédről, az éneklésről, ennek artikulációs, akusztikai és percepció vonatkozásairól, ami miatt minden, a tágabb értelemben vett beszédtudomány iránt érdeklődő olvasónak szívvel ajánlhatom.

Markó Alexandra

A szoprán énekes feladata – különösen magas alapfrekvenciákon – meglehetősen kihívásokkal teli, és igaz ez két okból is. Egyfelől azért, mert a magas alapfrekvenciák (magas zenei hangok) produkciója eleve nem mindennapi teljesítmény, melyre – mint a kötetben is röviden tárgyalni fogjuk – nem mindennapi artikulációs mozgások begyakorlása és alkalmazása teszi képessé az énekest. Másfelől pedig azért is, mert magas alapfrekvenciákon a hangmagasság produkálása és a magánhangzók ejtése az énekelt szövegben látszólag teljesen ellentmondanak egymásnak – legalábbis a beszédről szóló ismereteink alapján.

Az egyes magánhangzók minőségét és így a magánhangzók egymástól való elkülöníthetőségét a magánhangzók létrehozásakor az artikulációs szervek állását jellemző üregi rezonanciák, azaz a formánsok alakítják ki. A formánsok között kitéüntetett jelentőséggel az első és második formáns bír, ezek a formánsfrekvenciák ugyanis az egyes magánhangzókra jellemzőek, ráadásul az artikulációs szervek állásával viszonylag egyszerűnek tekinthető összefüggésben állnak. Az első formáns a nyelv függőleges helyzetével fordított arányban változik (a magasabb nyelvállás tehát alacsonyabb frekvenciaértékeket eredményez az első formánsra nézvést, így pl. az /i/ első formánsa alacsonyabb, mint a nála nyiltabb /a/ első formánsa). A második formáns pedig elsősorban a vízszintes nyelvhelyzet függvényében alakul (a hátul képzett, veláris hangzók, pl. az /u/ második formánsa alacsonyabb, mint az elől képzett, palatális magánhangzók, pl. az /i/ második formánsa). Ennek megfelelően a zárt magánhangzók, így például az /i/ vagy az /u/ első formánsa meglehetősen alacsony frekvenciaértékű – annál a frekvenciaértéknél legalábbis minden bizonnyal alacsonyabb, mint amilyen alapfrekvencián a magas hangosztályok énekesei, elsősorban a szopránok énekelnek, tehát amely alapfrekvenciákon a szopránok igen gyakran képeznek beszédhangokat éneklés közben.

Felmerül tehát a kérdés, mi történik a magánhangzók az adott hangzóminőség kialakításának szempontjából kritikus összetevőivel (az első vagy adott esetben akár a második formánssal) a magas alapfrekvenciájú éneklésben, tehát akkor, amikor az énekelt alapfrekvencia magasabb, mint a kérdéses formánsok átlagos (beszédben jellemző) frekvenciaértékei. Ha ugyanis az érintett formánsok egyszerűen csak „eltűnnének” a magánhangzó spektrumából (tehát az énekelt magánhangzókból csak a magasabb sorszámú formánsok rezonálnának), az azt eredményezné, hogy az egyes magánhangzó-minőségek is megszűnnének a magas alapfrekvenciákon, sőt a hang jelentősen veszítene észlelt hangerejéből is, miközben a produkált hangszín(ezet)ben is hirtelen és kellemetlen változást tapasztalnánk. Mivel azonban az énekesek általában jól érthetőnek tűnő, de feltétlenül nagy hangerejű, szép hangszínezetű hangot, nagy intenzitású magánhangzókat képeznek (ami pusztán a felsőbb sorszámú, ráadásul az artikulációs szervekkel nem vagy jóval kevésbé befolyásolható frekvenciájú rezonanciákkal

nehezen volna kivitelezhető), az alsóbb frekvenciaértékű formánsok egyszerű „elntünését” nem valószínűsíthetjük.

Mint azt a kötetben is látni fogjuk, az itt felvetett kérdés – természetesen – korántsem tekinthető egyszerűnek, ennek megfelelően a nemzetközi tudományos világban is tekintélyes szakirodalma van. Ezen irodalom tanúsága szerint a szoprán énekesek a magas alapfrekvenciájú éneklés során egészen különleges (és a mindennapi beszélők számára ismeretlen) hatékonysággal használják ki artikulációs szerveik és az artikulációs üreghendszer (elsősorban a garat és a szájüreg) rezonanciáit. E nem mindennapi akusztikus stratégiák célja egyfelől az, hogy a szoprán énekesek képesek legyenek kiénekelni a különlegesen magas alapfrekvenciákat (a különlegesen magas zenei hangmagasságokat), másfelől pedig az, hogy az így kiadott hang erőssége is elegendő legyen ahhoz, hogy az énekhang erősítés (mikrofon és hangfalak) nélkül is jól hallható legyen akár egy teljes zenekaron át – úgy, hogy közben a hangszínezetet is képes jól kontrollálni az énekes. Mivel azonban az ehhez a rendkívüli teljesítményhez használt rezonanciák egyúttal a magánhangzók kialakításáért is felelősek (lennének), és az énekesek az előadás során mindenképpen valamely nyelv magánhangzó- (és mássalhangzó-) készletéből építkező szövegeket énekelnek, felmerül a kérdés: mi történik az éneklésben ejtett magánhangzókakkal akkor, ha azok specifikus rezonanciáit (formánsait) az énekes inkább a szép és nagy hangerejű énekhangképzés szolgálatába állítja az egyes magánhangzók közti különbség megtartása helyett?

Az értekezésben erre a kérdésre keressük a választ mind az énekesek, azaz a produkció és az akusztikai megvalósítás, mind pedig a hallgatók, azaz az észlelés, a percepció oldaláról. Reményeink szerint e kérdés körülményeinek révén nem csak az énekhang-képzésről szerzünk új tudományos ismereteket, hanem az emberi beszédprodukció és a beszédpercepció működéseit is olyan új oldaláról ismerhetjük meg, melyre a mindennapi beszéd vizsgálata révén nem volna módunk.

A kötet *1. fejezete* a beszéd- és énekhang produkciójához szükséges beszédképző szervek és a beszéd képzésében játszott szerepük áttekintésével kezdődik, melyet előbb a magánhangzók, majd az énekelt magánhangzók már eddig feltárt artikulációs, akusztikai sajátosságainak, valamint percepciós jellemzőinek leírása követ. Az értekezés első fejezetének végén újabb megválaszolásra váró kérdéseket vetünk fel az éneklés produkciós és percepciós vetületét illetően, különös tekintettel a magyar magánhangzók megvalósítására az éneklésben. A kötet második részében a *2. fejezettől* a *7. fejezetig* a kérdések megválaszolásához lefolytatott saját kísérleteinket és azok eredményeit ismertetjük. Végezetül, az értekezés befejezéseként a *8. fejezetben* a kísérleti eredmények összefoglalásával az értekezésben bemutatott kutatómunka új eredményeit és következtetéseit vesszük sorra, megkísérlünk reflektálni eredményeink gyakorlati felhasználhatóságára, valamint felvetünk néhány további, még megválaszolásra váró kérdést, új vizsgálati irányt is.

A kötet létrejöttében nyújtott segítségéért számos volt tanáromnak, kollégámnak, barátomnak tartozom hálával. Köszönettel tartozom Gósy Máriának és Surányi Baláznak, akik segítették a doktori kutatásaim megvalósítását. Köszönöm Bárkányi Zsuzsannának, Bóna

Juditnak és Vicsi Klárának az értekezéssel kapcsolatos tanácsaikat, kritikai észrevételeiket. Köszönöm Gráci Tekla Etelkának, Mády Katalinnak és Vincze Miklósnak a rengeteg szakmai segítséget. Köszönöm Bódi Zsófia szoprán énekesnőnek, aki évek óta felbecsülhetetlen szakmai támogatást nyújt a kutatásaimhoz. Köszönöm az adatközlőimnek, hogy időt szántak a kísérleteimre, és külön köszönöm az énekeseknek a beszélgetéseinket, melyeken keresztül betekintést engedtek szakmájuk rejtelseibe. I am very grateful for all the invaluable opportunities I got from Sten Ternström and the KTH Royal Institute of Technology. I am also grateful to Svante Granqvist for his advice, support, and for his charity work of developing the freeware DeCap software I used for manual inverse filtering. And of course, my special thanks go to Johan Sundberg for his friendship and for all the constructive criticism and support I got from him. Köszönöm Markó Alexandrának, hogy a szakdolgozatom születésétől fogva témavezetőként támogat, a rengeteg szakmai és baráti segítséget, a lehetőségeket és főként – a bizalmat. Köszönöm, hogy a jelen könyv megjelenését lehetővé tette és gondos szerkesztői munkájával segítette. Végezetül pedig köszönet illeti a barátaimat is, és mindenkit, akivel az elmúlt tíz évben találkoztam. Valamint köszönettel tartozom Káldi Tamásnak is számos dologért, de legfőképpen azért, hogy töretlenül támogat.

1. Általános bevezetés

1.1. A magánhangzók produkciója és percepciója

1.1.1. A beszédkeltés folyamata, különös tekintettel a magánhangzók produkciójára

A beszédkeltés főbb szervei a tüdő, a légcső, a gége és az abban található hangszalagok, a garat, a szájüreg, valamint az orrüreg. A hangszalagoktól (vagyis a gégetől) kezdve a száj- és az orrüregig bezárólag ezek a szervek együttesen alkotják a szupraglottális üregrendszert, melyet átfogóan artikulációs csatornának, toldalékcsőnek vagy ritkábban (a *vocal tract* terminus magyarításával) vokális traktusnak is nevezünk. Ezzel párhuzamosan a gége, még pontosabban a hangrés alatti üregrendszer általánosan használt neve átfogóan szubglottális üregrendszer (GÓSY 2004).

A beszédkeltés mechanizmusát FANT (1960) úttörő munkája óta hagyományosan lineáris, egy forrásból és szűrőből álló rendszerben képzeljük el (source-filter theory, vö. még pl. GÓSY 2004; VICSÍ 2010b). Ebben a modellben a hangadás alapja a forrás (valamilyen periodikus vagy nem periodikus, azaz zörejszerű, szabálytalan rezgés), mely a szűrőn átjutva (az ott ért módosítások után) a szájon át kisugárzik – ez adja a hallható beszédet. A linearitás itt arra utal, hogy a beszédképzés folyamatában a hangadás alapját képező levegő, valamint a létrejövő hanghullámok a hangforrás felől a beszédképző szervek rendszerén át csak egy irányba terjednek. Ez, bár a valóságban nem pontosan így van, de a modell számára mégis jó közelítést ad az artikulációs csatorna működéséről, és egyúttal (általában) elegendő a létrejövő akusztikai jel alapvető jellemzőinek megértéséhez. A forrás és szűrő elnevezések arra utalnak, hogy a hangadást, beszédkeltést valamilyen hangforrás (a magánhangzók esetében ez a zöngé) létrehozásával, majd pedig a létrehozott hang módosításával, szűrésével modellezzük (lásd VICSÍ 2010b).

A magánhangzók képzésekor a beszédkeltés során a tüdőből kiáramló levegő megrezgetti a gégeben található hangszalagokat (ez lesz tehát a forrás), az így megrezgetett levegő (más néven zöngé) pedig a toldalékcsővön keresztülhaladva különböző akusztikai módosulásokon esik át (a toldalékcső, tehát a szűrő hatására), mígnem végül a szájnyílást elhagyva elnyeri végleges formáját, és magánhangzóvá alakul. Az itt következőkben ezt a folyamatot, azaz a beszédkeltés közismert fiziológiai és akusztikai folyamatait tekintjük át röviden a forrás

és a szűrő, tehát a zönge és a toldalékcso által okozott akusztikai változások elválasztásával, különös tekintettel a magánhangzóképzésre Vicsi (2010b), valamint Seikel és munkatársai (2010) alapján.

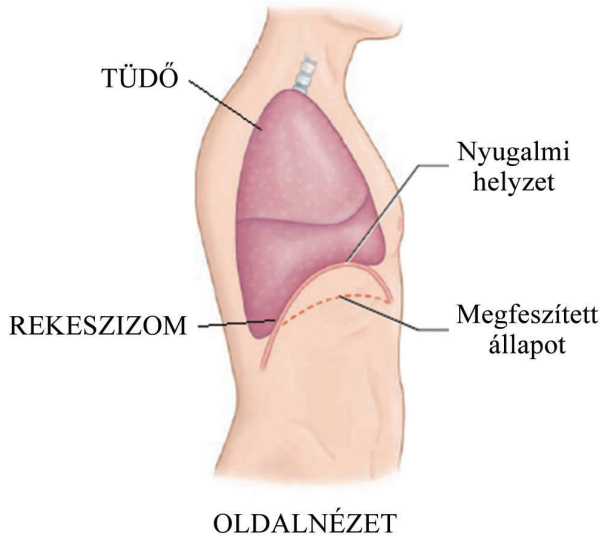
1.1.1.1. A ZÖNGE (AZAZ A FORRÁS) KELETKEZÉSE, AZ ALAPFREKVENCIA

A beszédkeltés a belégzéssel kezdődik, mely során a tüdő megtelik levegővel. A belégzésben a tüdő szerepe passzív, tehát a levegőnek a tüdőbe jutása nem a tüdő öngerjesztett mozgásának eredménye. A tüdő a mellkasban helyezkedik el, a bal és jobb tüdő a mellkasba a légcsőhöz alulról csatlakozva lóg. A tüdőhöz közvetlenül nem kapcsolódnak izmok, melyekkel mozgatható lenne, közvetve azonban, a kettős falú mellhártya révén mégis szoros összeköttetésben áll a bordákkal, valamint a tüdőt alulról megtámasztó rekeszizmokkal is. A tüdőt a mellkashoz és a rekeszizomhoz is negatív nyomás, azaz egyfajta szívó hatás rögzíti, mely a mellhártya két rétege között áll fenn. A mellhártya egyik rétege a tüdőhöz, míg a másik a mellkashoz kapcsolódik. A szívó hatás ezt a két réteget húzza egymáshoz. A mellhártyának köszönhetően tehát mozgatható, egészen pontosan kifeszíthető a tüdő, ugyanis a mellkas és a rekeszizom mozgását a mellhártya által (passzívan) a tüdő is követi. Mivel a tüdő fentről és oldalról a bordákhoz, alulról pedig a rekeszizomhoz van rögzítve, azok ellentétes irányú mozgása a tüdő rugalmas lebenyeit kifeszíti, ezzel megnövelve annak térfogatát. Belégzéskor éppen ez történik: a rekeszizom megfeszül, lefelé húzódik, a mellkas pedig a bordaközi izmok megfeszítése révén megemelkedik (*I. ábra*). Így belégzéskor a tüdő térfogata megnő.

A belégzés, azaz a levegőnek a tüdőbe juttatása ugyanazon az elven alapszik, mint a fecskendő működése: mivel a gázok mindig igyekeznek kitölteni a rendelkezésükre álló teret, ha a tér kitágul, a benne lévő gáz is tágulni kezd, ami pedig nyomáscsökkenést, azaz egyfajta szívó hatást eredményez. A tüdő a felnőttek esetében nyugalmi állapotban is ki van feszítve a mellkasban, így ebben az állapotában is viszonylag sok „maradványlevegőt” tartalmaz. Gyermekek esetében azonban a tüdő nyugalmi állapotában tapasztalható nagyobb fokú ernyedsége miatt ez a levegőmennyiség jóval kisebb. Amikor belégzéskor a bordák és a rekeszizom működése révén a tüdő kitágul, az abban lévő levegő is tágulni, a tüdőben lévő nyomás pedig csökkenni kezd. Az alacsony nyomás kiegyenlítődesre törekszik, és mivel a tüdő a hangszalagok nyitott állásában nem képez teljesen zárt rendszert, így a belégzéskor fellépő nyomáscsökkenés kiegyenlítődes. A kiegyenlítődes során a nyílásnál, azaz a hangszalagoknál lévő (a száj- vagy orrnyíláson keresztül a toldalékcsoého jutott) levegő a tüdőbe áramlik.

A légzés funkcióját tekintve lehet élettani légzés, mely során pusztán az életben maradáshoz szükséges gázcseré zajlik, valamint beszédlégzés is, mely során a belégzést hangadás, beszédkeltés követi. Az előbbi esetben a kilégzés alapvetően passzív folyamat, melyet a megfeszített bordaközi izmok és a rekeszizom elernyesztése, a mellkas lesüllyedése idéz elő. A beszédképzés során azonban a levegő kiengedése fokozatos (és a belégzési szakasznál

hosszabb időbeli lefutású), amit a beszélő aktívan is befolyásol a belső bordaközi izmok, valamint a hasfal izmainak munkája révén. A beszédkeltés közbeni aktív kilégzés tehát biztosítja a levegőtartalék folyamatos adagolását.

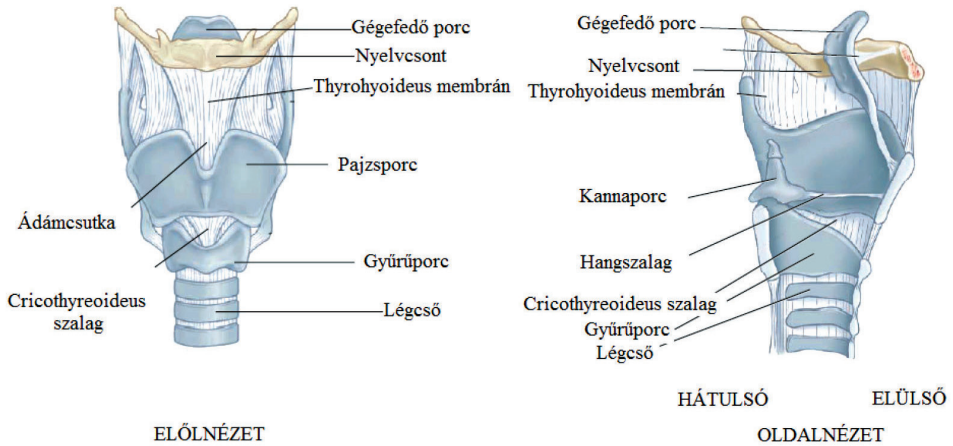


1. ábra: A rekeszizom működése (SEIKEL et al. 2010: 105 alapján)

A beszédképzés során a belégzés következtében a tüdőben felgyűlt levegő kipréselődik a tüdőből, és a légcsővön keresztül a gégebe jut. A légcső C alakú porcokból álló hajlékony, mégis kemény falú cső, így a nyak szabadon mozgatható, miközben a levegő energiája közel veszteségmentesen áramolhat át a légcsővön (azaz a kiáramlás alatt a levegő sebessége nem csökken jelentősen). A gége a légcső tetején helyezkedik el, porcokból, izmokból és szalagokból áll. Vázát öt porc adja: a gyűrűporc, a pajzsporc, a kannaporcok (egy pár) és a gégefedő porca (lásd 2. ábra). Ezeket a porcokat számos izom és szalag szövi át, így a gége igen változatos mozgásokra képes. A gégeben húzódnak a hangszalagok, melyek elülső oldala a pajzsporchoz kapcsolódik, hátulsó oldala pedig a kannaporcokhoz ízesül. A hangszalagok előlről mereven vannak rögzítve, hátulról azonban a kannaporcok mozgása révén változtatható a beállításuk – így a két hangszalag közelíthető (addukció) és távolítható (abdukció).

A gége a beszédképzésen kívül számos élettani funkcióval bír. A hangszalagok közti rés, a hangrés vagy glottisz bezárása egyrészt megakadályozza az idegen anyagoknak a tüdőbe jutását, másrészt képessé tesz minket arra, hogy elzárjuk a kifelé áramló levegő útját, azaz a tüdőnkben tartjuk a levegőt. A levegő visszatartása szükséges ahhoz, hogy tartalékolni tudjunk az életben maradáshoz szükséges gázokat (így lehetünk képesek például a víz alatt úszni), de ahhoz is, hogy fixáljuk, megmerevítsük a mellkasunkat, és ezzel nagyobb erő kifejtésre legyünk képesek. A gége élettani szempontból elsődleges szerepe a tüdőhöz vezető út

védelme és szükség szerinti gyors lezárása. Jól mutatja ezt, hogy a hangrés zárásához három pár adduktor izom specializálódott, míg a hangrés nyitását mindössze egy pár abduktor izom végzi. A glottiszt védő szereppel bír a felette található gégefedő vagy epiglottisz is, egy levél alakú porc, mely nyeléskor a hangrészt lezárva megakadályozza a szájartalom tüdőbe jutását.

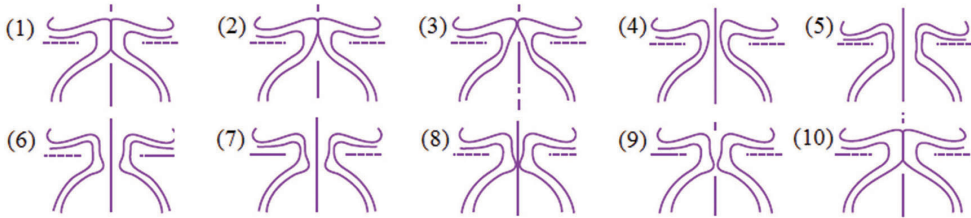


2. ábra: A gége felépítése (SEIKEL et al. 2010: 167 alapján)

A normál módú beszédképzés során a hangszalagok számos beállítása lehetséges. A hangszalagok e lehetséges beállításai közül azonban itt most csak a (magyar) magánhangzók képzésének szempontjából legfontosabb állást, a zöngképzés során használt zöngéállást tárgyaljuk. A zöngés hangok, tehát például az összes magyar magánhangzó képzésekor a hangszalagok a hangadás kezdetekor, a zöngeciklus létrehozásának kezdetén zárt pozícióban állnak: a két hangszalag közti rés, a glottisz vagy hangrés teljesen zárva van (lásd 3. ábra, 1. panel). A zárt hangszalagok ekkor akadályt képeznek a kiáramló levegő útjában, ezért a levegő felgyűlik a zárt hangszalagok alatt. Ha a felgyűlt levegő mennyisége és az így létrejött nyomásnövekedés megfelelő mértékű (illetve a szubglottális nyomás kellő mértékben nagyobb, mint a szupraglottális), a levegő szétfeszíti a hangszalagokat, és tovább áramlik a száj- vagy az ornyílás felé.

A hangszalagok távolodása alulról felfelé hullámként terjed (lásd 3. ábra, 2–5. panel). Mivel a hangszalagok szűkületet képeznek a gégében, a köztük átáramló levegő a légső és a toldalécső felsőbb szakaszához képest is gyorsabban halad. Ez Bernoulli törvénye értelmében azt vonja magával, hogy amikor a levegő áthalad az éppen csak szétfeszített hangrésen, a hangszalagok közötti nyomás lecsökken, ez pedig szívó hatást fejt ki a légjárat falaira, azaz a hangszalagokra. A szívó hatás következtében a szétfeszítés után a hangszalagok újra összetapadnak, így a zöngképzés egy új ciklusa kezdődhet meg. A hangszalagok közeledése és újbóli záródása a Bernoulli-effektus hatására szintén alulról felfelé terjedő állapot (lásd 3. ábra, 6–10. panel). A hangszalagok ciklikus nyitódása és összezáródása a **zöngképzés**

vagy más néven **fonáció**. A folyamat addig tart, amíg elegendő levegő áramlik ki a tüdőből ahhoz, hogy szétfeszítse az összezárt hangszalagokat, illetve amíg az így keletkezett résen átáramló levegő ismét képes összeszívni azokat (a kellően nagy sebesség, azaz megfelelően alacsony nyomás és a szub- és szupraglottális terek között fennálló megfelelő nyomáskülönbség segítségével).



3. ábra: A zöngéképzés egy ciklusa (FLEMMING 2005 alapján)

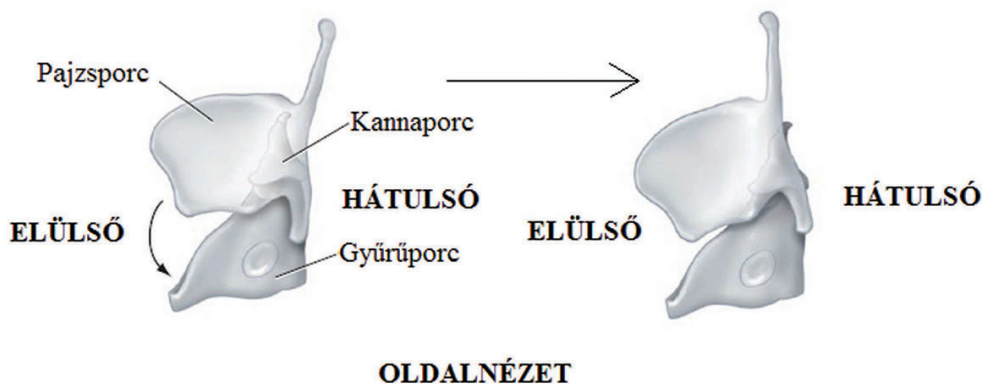
A hangszalagok nyitódása és záródása következtében a tüdőből még egyenletesen áramló levegő a géget már kváziperiodikus löketekben hagyja el, így az hangnyomás-ingadozást okoz, mely az idő függvényében (a hangszalagok ciklikus összezáródási fázisával összhangban) tehát szintén kváziperiodikus lesz. A zöngéképzés vagy fonáció következtében létrejött nyomásingadozást, azaz hangot nevezünk **zöngének**. A zöngé karakterisztikus jellemzője az **alapfrekvenciája**, azaz a zöngéképzés során az egy időegység alatt lezajlott ciklusok száma. Ez adja a zöngés hangoknak, például a magánhangzóknak az emberi fül számára is érzékelhető alaphangját, hangmagasságát – míg a rezgés amplitúdója, „nagyága” a hangerősségérzetet befolyásolja. Az alapfrekvencia jelölése általánosan f_0 , (fizikai, lineáris) mértékegysége a Hertz (Hz).

Bár úgy tűnik, hogy az alapfrekvenciára áttételesen számos tényező hatással van, közvetlenül gyakorlatilag mindegyik a hangszalagok méretét és egységnyi hosszra jutó tömegét befolyásolja, ez tehát az a jellemző, amely az alapfrekvenciát meghatározza. Az alapfrekvenciát meghatározó tényező lehet az életkor és a nem (hiszen a hangszalagok hossza és tömege általánosan különbözik a gyermek, a felnőtt női és a felnőtt férfi beszélők esetében), a hangszalagok feszítettsége, a gége vertikális és horizontális helyzete (ez ugyanis szintén hat a hangszalagok feszítettségre) vagy az ejtett beszédhang minősége (mely szintén összefüggést mutat a függőleges gégehellyeztetel, így a hangszalagok feszítettségével).

Noha a hangszalagok mozgását és annak sebességét elsősorban és közvetlenül csak a hangszalagok méretei, valamint az aerodinamikai hatások befolyásolják, ezekre közvetve az egyes gégeizmok működése is fontos hatással van. Az alapfrekvencia, azaz az alaphangmagasság emelését a hangszalagok megfeszítésével, megnyújtásával érjük el, így tudjuk ugyanis csökkenteni a hangszalagok egységnyi hosszra eső tömegét. Az alapfrekvencia csökkenését ezzel ellentétesen a hangszalagok ellazítása okozza. A hangmagasság emeléséhez a gyűrűporcot

és a pajzsporcot összekötő cricothyroideus izom, valamint a pajzsporcot és a kannaporcokat összekötő belső thyrovocalis izom megfeszítésére van szükség. Az előbbi megfeszítésének hatására a pajzsporc előrebillen, a hangszalagok megnyúlnak, így a zöngéképzés során létrehozott hang alulfrekvenciája is megemelkedik (lásd 4. ábra). Az utóbbi megfeszítése (antagonistaként) szintén ezt a folyamatot segíti.

Az alulfrekvencia csökkentéséhez az ellenkező irányú manipuláció szükséges, tehát a hangszalagok ellazítása. Ezt a pajzsporcot és a kannaporcokat összekötő külső thyromuscularis izom megfeszítésével érjük el, az ugyanis közelebb húzza egymáshoz a pajzsporcot és a kannaporcokat, így rövidíti, lazítja a hangszalagokat. Mivel a magasabb alulfrekvencia képzéséhez megfeszített hangszalagokat nehezebb mozgásba (azaz rezgésbe) hozni, az alulfrekvenciaemeléshez a szubglottális nyomás emelése is szükséges. Ebből következően a magasabb alulfrekvenciájú hangok ejtését mindig magasabb szubglottális nyomás, azaz magasabb hangnyomásszint, nagyobb hangerősség is kíséri.



4. ábra: A gégeporcok mozgásai az alulfrekvencia emeléséhez: az alulfrekvencia emeléséhez a pajzsporc előrebillen, így jobban megfeszíti az előlről a pajzsporchoz, hátulról pedig a kannaporcokhoz csatlakozó hangszalagokat
(SEIKEL et al. 2010: 185 alapján)

A gégében keletkező zöngé tehát meghatározza a keletkező zöngés beszédhang alulfrekvenciáját, azaz annak észlelhető hangmagasságát. Ám a zöngé **komplex hang**, azaz nemcsak a hangszalagok rezgésszámára jellemző alulfrekvenciát, hanem annak egész számú többszöröseit, azaz a **felharmonikusokat**, **felhangokat** is tartalmazza.¹ Az alulfrekvencia és a felharmonikusok jelölése többféle lehet. Ezek közül a jelen kötetben az egyszerű átláthatóság kedvéért (valamint azért, hogy a jelölés felidézze a felhangoknak az alulfrekvenciával való

¹ A felharmonikusok létrejöttének oka egyébként az, hogy a felnyíló glottison átáramló levegőmennyiség áramlási sebessége valójában a levegő tehetetlensége miatt nem szinuszos, hanem elferdült jelalakban írható le. Ez az elferdült jelalak viszont már egy komplex hang hullámformája: nemcsak a rezgés alulfrekvenciáját, hanem annak egész számú többszöröseit is tartalmazza (vö. Vicsi 2010b).

kapcsolatát) az f_0 , valamint a $2f_0$, $3f_0$, $4f_0$ stb. notációt alkalmazzuk, rendre az alaphfrekvencia és a felharmonikusok jelölésére, a legalacsonyabb frekvenciájú összetevőtől (az alaphfrekvenciától) a magasabbak felé haladva. A zöngében a felharmonikusok amplitúdója egyenletes, 12 dB/oktáv csökkenéssel van jelen. Ez jellemzi tehát a zöngé spektrumát, azaz a zöngé frekvencia-összetevőinek amplitúdóját a frekvencia függvényében – a folyamat e pontján még elvben (a lineáris forrás-szűrő modell szerint) minden esetben, az ejteni szándékozott magánhangzó minőségétől függetlenül. Ahhoz, hogy a zöngéből különböző magánhangzó-szerű hangok keletkezzenek, szükséges, hogy a zöngé keresztülhaladjon a toldalékcsővön, mely akusztikus szűrőként változtatja meg a zöngének az ember számára hangszinként érzékelhető spektrumát, azaz a zöngé egyes felhangjainak intenzitását.

1.1.1.2. A TOLDALÉKCSŐ MINT AKUSZTIKUS SZŪRŐ, A FORMÁNSOK

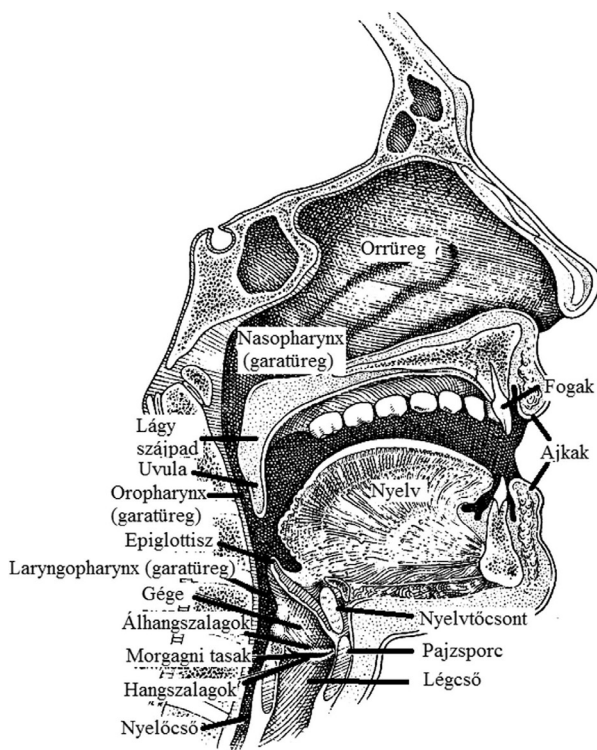
A toldalékcső főbb részei a garatüreg (pharynx), a szájüreg (cavum oris) és az orrüreg (cavum nasi). Ezekben az üregekben passzív és aktív artikulációs szervek találhatók, melyek a toldalékcső alakját módosítva változtatják annak sajátfrekvenciáit, azaz a toldalékcső szűrőtulajdonságait (lásd 5. ábra). A jelen fejezetben előbb az artikulációs szerveket és főbb beszédbeli funkcióikat tekintjük át – különös tekintettel a magánhangzók képzésében részt vevő szervekre –, majd az üregi rezonanciák keletkezését tárgyaljuk.

Az aktívan mozgatható artikulációs szervek közül a legnagyobb és talán a legfontosabb a nyelv és az állkapocs, melyek működése lévén a szájüregben kisebb vagy nagyobb szűkület, illetve rést vagy zárat képzünk a nyelv és a szájpad között az egyes beszédhangok ejtése során. A nyelv mozgásai igen változatosak: a nyelv a külső izmok munkája révén vízszintesen és függőlegesen is mozdítható szerv (a külső izmok határozzák meg tehát a nyelv helyzetét a szájüregben), míg a belső izmok segítségével alakítható a nyelv alak (domborítható, egyenesíthető a nyelvhat stb.). A nyelv mozgása révén egyfelől befolyásolja a garat méretét, másfelől pedig a nyelvet a nyelvcsonttal összekötő izmok révén a gége helyzetét is. A nyelv főbb részei a nyelv csúcsa vagy nyelvhegy (apex/cacumen), a nyelvperem vagy nyelvparkány (corona), a nyelvhat (dorsum), valamint a nyelvgyök (radix). A nyelvhatat szokás további három részre is tagolni, ezek az elülső (predorzális), a középső (mediodorzális) és a hátsó (posztdorzális) részek. Bár a nyelv mozgása (elsősorban a függőleges helyzetének módosítása) és az állkapocs nyitáshozogének változtatása szorosan összefügg, az egyik helyzete nem törvényszerűen determinálja a másikat, hiszen bizonyos mértékig külön-külön is mozgatható artikulátorokról van szó – erre a későbbiekben még a magyar magánhangzórendszer taglalásánál, valamint az éneklés artikulációjának és akusztikai szerkezetének bemutatásánál is utalni fogunk.

Az aktív, mozgatható artikulátorok közé tartoznak még az ajkak, a garatüreg fala, valamint a gyakorlatilag csak izmokból álló lágy szájpad (a szájpadlás hátulsó része, a velum palatinum) és annak zárata, a nyelvcsap vagy uvula. Az ajkak főbb mozgásai a távolodás vagy a közeledés (egészen a teljes zárképzésig), valamint az ajkak kerekítése és széthúzása, illetve

előrecsücsörítése. A lágy szájpad és az uvula legfőbb funkciója, hogy az orrüreg és szájüreg közti átjáró lezárásán vagy megnyitásán keresztül elzárja vagy megnyitja a levegő útját az orrüreg felé, így rendre az orális és nazális hangminőségek létrehozásáért felelős.

A passzív vagy nem mozgatható artikulátorok közé tartoznak a fogak, a fogmeder és a kemény szájpad (mely a szájpadlás elülső része, palatum durum). Ezek azok a szervek, melyekhez az aktív, mobilis artikulátorok közelednek a beszédképzés során (vö. 5. ábra).

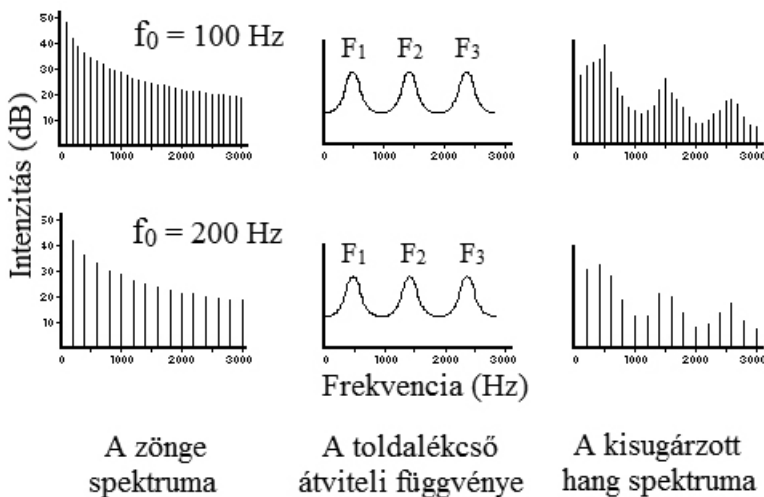


5. ábra: A toldalékcső részei (SUNDBERG 1977: 83 alapján)

A magánhangzók ejtésékor a nyelv jellemzően nem érintkezik a szájpadalattal, csak közelít hozzá, a levegő a toldalékcsőben akadálymentesen áramlik (a rés vagy zár létrehozása elsősorban a mássalhangzók képzésére jellemző). Megjegyzendő mégis, hogy a magánhangzók képzésében is szerepe van annak, hogy a nyelv egyes részei érintkeznek a szájüreg más (immobilis) artikulációs szerveivel, mert ezek a szenzoros visszajelzések artikulációs sarokpontokat jelentve feltehetőleg hozzájárulnak ahhoz, hogy a beszélő az egyes hangok ejtéséhez szükséges artikulációs beállításokat pontosan eltalálja (STEVENS 1998). Az orális magánhangzók – így például a magyar magánhangzók – képzése szempontjából a legfontosabb artikulációs szerveknek a szájüregben létrehozott szűkület méretét befolyásoló nyelvet és állkapcsot, valamint az ajkakat tekinthetjük.

A toldalékcső mint üregrendszer más üregekhez hasonlóan rendelkezik bizonyos sajátfrekvenciákkal, melyekkel gerjesztve rezonálni kezd. Ezek a frekvenciák együttesen adják meg a toldalékcső átviteli karakterisztikáját, azaz azt a függvényt, mely a sajátrezonanciákat „összesíti”. Amint azt már említettük, a zöngé mind az alapfrekvenciáján, mind annak egész számú többszörösein tartalmaz frekvenciaösszetevőket, felharmonikusokat. Ennek köszönhetően a mindennapi (nagyon átlagosan körülbelül 100–400 Hz közti alapfrekvenciájú, vö. Gósy 2004) beszéd során a zöngé akusztikai szerkezete a toldalékcsőre jellemző frekvenciahelyeken akár több felharmonikkal is rendelkezik. Így tehát ha van fonáció, azaz zöngét képezünk, gerjesztjük a toldalékcövet, mely az alakjának megfelelő frekvenciákon a zöngével együtt rezonálni kezd. Ennek a rezonanciának az eredményeképpen a zöngé spektruma megváltozik, azaz a szájnál kisugárzott hang spektruma eltérő lesz annak a hangnak a spektrumától, amely a géégében létrejött. A toldalékcső ugyanis a rezonanciák révén megszűri a zöngéhangot: a toldalékcső felerősíti a zöngében a sajátfrekvenciáinak frekvenciasávjába eső felhangokat vagy felhangnyalábokat, és csillapítja az azoktól távolabb eső frekvenciaösszetevőket (azaz csökkenti azok amplitúdóját). Ezt a folyamatot szemléltetjük a 6. ábrán.

A beszéd szempontjából a toldalékcső rezonanciahelyeinek vagy sajátfrekvenciáinak kiemelt jelentősége van, hiszen ezek határozzák meg a kisugárzott hang specifikus hangszínét, azaz az egyes beszédhangokat, a beszédhangok minőségét. A toldalékcső ezen állóhullámú rezonanciáinak elnevezése a beszédtudományban **formáns**, jele pedig F_1 , F_2 , F_3 stb. a legalsócsnyabb frekvenciaértékű rezonanciától kezdve a sorszámozást.



6. ábra: A zöngés magánhangzók létrehozásának folyamata a forrás-szűrő modellben egy mélyebb (felső panel) és egy magasabb (alsó panel) alapfrekvenciájú hang esetében (<http://www.haskins.yale.edu/featured/heads/mmsp/acoustic.html> alapján)

Mivel a felharmonikusok sűrűsége az alapfrekvenciától függ, ezért elmondható, hogy minél mélyebb a beszédhang, annál több frekvenciakomponense esik a toldalékcső sajátrezonanciáinak frekvenciasávjába, és annál több felharmonikusa vesz részt a beszédjel akusztikai kimenetében található frekvenciamaximumok létrehozásában. Ezt az összefüggést figyelhetjük meg a 6. ábra alsó és felső panelének összevetésével: az alsó panelen egy magasabb, 200 Hz alapfrekvenciájú, míg a felső panelen egy alacsonyabb, 100 Hz alapfrekvenciájú hang keletkezését látjuk. A formánsok és a toldalékcső sajátfrekvenciáinak az imént megfogalmazott összefüggése a magas alapfrekvenciájú éneklésben kiemelt jelentőségű (így, mint látni fogjuk, a jelen értekezésben is kiemelt figyelmet kap). A magas alapfrekvencia mint akusztikus tulajdonság ugyanis problematikusá teszi mind a beszédhangok formánsainak detektálását, mind pedig a magas alapfrekvencián ejtett beszédhangok percepciójának, észlelésének értelmezését. Ezekről a későbbiekben még részletesebben is szót ejtünk.

Annak érdekében, hogy a különböző beszédhangok létrejöttét és a beszédhangokra jellemző frekvenciák kialakulását megérthessük, a beszédkutatók számos modellt alkottak. Ezekben közös, hogy a toldalékcsővet egy vagy több (ez esetben különböző keresztmetszeti területű) csőhöz hasonlítják, ezzel ugyanis jó közelítés adható a toldalékcsőnek az egyes magánhangzók ejtésekor jellemző alakjáról. Így a magánhangzókra jellemző frekvenciahelyeket, azaz formánsokat a valóságban igen bonyolult szerkezetű toldalékcső helyett egy annál jóval egyszerűbb szerkezetű csőből vagy csőrendszerből vezethetjük le. Ezen elméletek a toldalékcsővet általánosan egy az egyik végén (a gégénél) zárt, a másik végén (a szájnyílásnál) pedig nyitott csőnek tekintik. Az ilyen csőben a gerjesztés hatására ébredő állóhullámú rezonanciák, azaz a cső sajátfrekvenciái adott keresztmetszet mellett csak a cső hosszától függenek. A cső állóhullámú rezonanciái 17 cm hosszúság (azaz egy átlagos férfi toldalékcsőméret) esetén a peremfeltételek figyelembevételével (tehát figyelembe véve, hogy a cső egyik oldalán zárt, míg a másik oldalán nyitott rendszer) megközelítőleg 500, 1500 és 2500 Hz.

A **perturbációs elmélet** a toldalékcsővet egyetlen (például az iménti példában megjelölt sajátrezonanciákkal rendelkező) csőként képzeli el, melynek rezonanciáit a beszéd során a beszélő folyamatosan módosítja, mégpedig úgy, hogy a rezonanciák minimum- és maximumhelyeinél (azaz a maximális vagy minimális térfogatsebességeknél) szűkíti a toldalékcsővet. Az elmélet szerint a térfogatsebesség maximumhelyeinél okozott szűkület csökkenti az adott rezonancia értékét, míg a minimumhelyeknél okozott szűkület növeli azt (CHIBA-KAJIYAMA 1941; idézi FLEMMING 2005).

A **beszédképzés gerjesztett cső modellje** – vagy röviden az **akusztikus csőmodell** – azonban a magánhangzók ejtésekor létrehozott csőkonfigurációt nem egy, hanem két akusztikusan csatolt csővel modellálja, melyek hossza és keresztmetszete az egyes magánhangzók ejtésekor jellemző toldalékcsőformának megfelelően különböző lehet. E modell szerint az egyes magánhangzók eltérő formánszerkezetét a két csatolt cső a csőhosszoktól függő sajátrezonanciái, valamint a két cső csatolásának mértéke alakítja ki, ami a két cső keresztmetszeti területeitől, illetve azoknak az arányától függ (STEVENS 1998).

Amint azt láttuk tehát, a formánsok frekvenciáit és sáv szélességét a toldalékcső méretei, a toldalékcső hossza és alakja, vagyis az egyes szakaszokon jellemző keresztmetszeti területe határozza meg. Bár ezek a paraméterek alapvetően minden beszélő esetében fiziológiailag adottak, mégis bizonyos kereteken belül változtathatók, éppen ezért az üregrendszer sajátfrekvenciái nem állandók.

A toldalékcső hossza egy adott beszélő esetében egy adott életkorban csak kisebb mértékben befolyásolható, a változás a függőleges gégehelyzet módosításával (a gége lesüllyesztésével vagy felemelésével), illetve az ajkak csücsörítésével érhető el. A toldalékcső hosszát legnagyobb mértékben a fizikai érés folyamata változtatja meg, hiszen az általános testi növekedéssel együtt a toldalékcső is meghosszabbodik.

Ezzel szemben a toldalékcső keresztmetszeti területe (éppen a jelen alfejezet elején taglalt artikulációs szervek mozgásai révén) egy adott életkorban is viszonylag nagymértékben változtatható. Mivel pedig a toldalékcső alakjának változtatásával valójában annak formánsfrekvenciáit módosítjuk, elsősorban ez a toldalékcsőnek azon tulajdonsága, melyet a beszédképzés során is kihasználunk ahhoz, hogy egymástól eltérő spektrumú, azaz eltérő hangszínű, eltérő minőségű magánhangzókat képezzünk. Az egyes magánhangzók ejtése ilyen módon felfogható a rezonátorüreg egy adott artikulációs konfigurációra hangolásával, aminek következtében a toldalékcső sajátfrekvenciái adott, az egyes magánhangzókra jellemző értékeket vesznek fel. Ennek megfelelően a magánhangzókat artikulációsan és (az artikulációs működésekkel szoros összefüggésben) akusztikailag is jellemezhetjük. Ezt tárgyaljuk a következő alfejezetben.

Mielőtt még azonban áttérnénk a magánhangzók artikulációs és akusztikai jellemzésére, fontosnak tartjuk két fogalom rövid tisztázását. Az egyik fogalom a hangszín, valamint annak elkülönülése a hangszínezettől, a másik pedig a formáns.

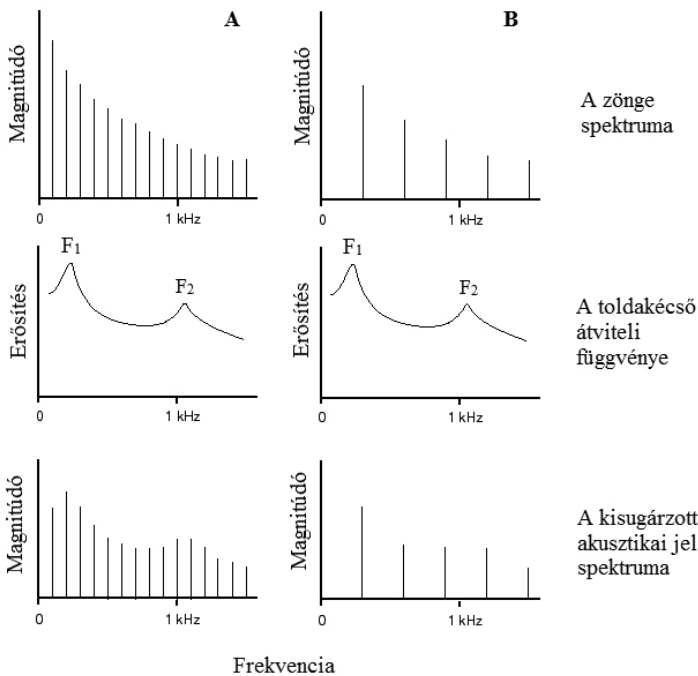
A hangoknak az ember számára érzékelhető minőségét **hangszínnel** nevezzük. A hangszín pszichoakusztikai fogalom, az ANSI (American National Standards Institute, 1960) meghatározása szerint az egyes hangok azon tulajdonsága, mely alapján a hallgató képes két ugyanolyan hangerejű és ugyanolyan alfrekvenciájú hangot különbözőkként felismerni. Bár a hangszín elsődlegesen a hang spektrumának függvénye, függ a hang hullámformájától, a hangnyomástól, valamint a hang időszerkezetétől is (American Standards Association, 1960). A hangszín tehát az, ami például elkülöníti egymástól a különböző hangszerek hangját – és mivel a hangszerek testének alakja vagy anyaga (szinte) sosem változik, ez a meghatározás a hangszerek világában elegendő is. Más a helyzet azonban az emberi beszéd esetében, itt ugyanis a hangszín az egyes beszélők eltérő egyéni hangszínére és az egyes magánhangzók egymástól eltérő hangszínére is utalhat. Éppen ezért Gósy (1999) a jelenségek megnevezésének egyértelműsítése céljából megkülönbözteti a **hangszín** és a **hangszínezet** fogalmakat, az előbbit a különböző magánhangzó-minőségekre, az utóbbit pedig az egyéni beszédjellemzőkre értve. A jelen értekezésben ezt a terminológiai elválasztást követjük, így a magánhangzókat egymástól elkülönítő akusztikai-auditív sajátosságokra a hangszín terminussal, a beszélő hangjának vagy az éneklésnek a sajátos akusztikai „színezetére” pedig a hangszínezet terminussal utalunk.

A második, az értekezés kérdésfelvetései szempontjából alapvetően fontos fogalom a **formáns**, ez ugyanis – bár nagyon gyakran használt és alapvetőnek tekintett – mégsem egyértelmű terminusa a beszéddel foglalkozó tudományoknak. A formáns első definícióját Gunnar FANT (1960) adta, e szerint a formánsok a hangspektrum csúcsai, vagyis a hangnyomás-idő függvény spektrumának energiamaximumai. Emellett azonban FANT (1960) megemlítette azt is, hogy a kimeneti energiamaximumok frekvenciahelyeihez közel eső, a toldalécső átvitelére (gain function 'erősítés függvény' $|T(f)|$) jellemző rezonanciafrekvenciákat a formánsoktól elkülönülten kell kezelni, annak ellenére, hogy ezek valóban sokszor egybesnek a kimenetben mérhető csúcsokkal. FANT (1960) tehát konceptuálisan elkülönítette a toldalécsőre jellemző sajátfrekvenciákat (ezeket *rezonanciáknak* nevezte), valamint a kimenetben mérhető energiacsúcsokat (melyekre a *formáns* elnevezést használta). A szakirodalom azonban nem egységes e hagyomány követésében, a *formáns* kifejezést ugyanis sok kutató minden esetben (tehát alapprofreciától függetlenül) az akusztikai kimenetben mérhető csúcsokra érti (lásd az Amerikai Akusztikai Társaság meghatározása (ANSI 1994), valamint pl. BENADE 1976; GÓSY 2004; NÉMETH–OLASZY szerk. 2010), míg mások a toldalécső sajátrezonanciáinak megjelölésére használják (lásd pl. STEVENS–HOUSE 1961; SUNDBERG 1975). Az előbbi meghatározás értelmében definiált formánsok (azaz a spektrális energiamaximumok) alacsonyabb alapprofreciák esetében általában egybesnek az utóbbi értelemben definiált formánsokkal (azaz az üregi sajátrezonanciákkal), de magasabb alapprofreciák esetében a 6. ábra kapcsán taglalt összefüggés miatt nem. Mivel azonban az első meghatározás értelmében definiált formánsok frekvenciaértéke alapprofreciára függő, ezért az e definíció szerint meghatározott *formáns* elnevezés magasabb alapprofreciákon már nem értendő (és nem is érthető) a toldalécsőre jellemző sajátrezonanciákra, hanem kizárólag az akusztikai kimenetben mérhető spektrális energiamaximumokra. Ezzel együtt viszont, ha a *formáns* terminust a kimeneti energiamaximumok megnevezésére alkalmazzuk, az artikulációval szoros összefüggést mutató rezonanciákra magas alapprofrecián már nem áll rendelkezésünkre elnevezés, azok tehát az első meghatározás értelmében magas alapprofrecián gyakorlatilag értelmezhetetlenné válnak. Mindemellett a szakirodalomban azt tapasztaljuk, hogy a formánsokat az első definíció, tehát az akusztikus kimenet szerint meghatározó vizsgálatok is gyakran az artikulációra utaló tulajdonságokként, illetve rezonanciákként tárgyalják. Teszik ezt annak ellenére, hogy bár a rezonanciák alacsonyabb alapprofreciákon viszonylag nagy pontossággal kikövetkeztethetők a kimenetben mérhető spektrális energiamaximumok frekvenciaértékéből, magasabb alapprofreciákon azonban nem, és ezzel együtt a spektrális csúcsokból magas alapprofreciákon a beszélő artikulációs (és akusztikai) stratégiáira sem következtethetünk. Mindezt figyelembe véve úgy látjuk helyesnek, ha a jelen értekezésben – melyben a magas alapprofreciák kérdése igen fontos lesz – a fent megfogalmazottak közül a második formánsdefiníciót tekintjük iránymutatónak.

A jelen kötetben formánsnak elsősorban nem az akusztikai kimenetben mérhető spektrális energiamaximumokat, hanem az egyes hangzók ejtésekor az artikulációs konfigurációra jellemző átviteli függvény csúcsait, azaz a toldalécső sajátfrekvenciáit, rezonanciáit tekintjük.

Az értekezés azon pontjain, ahol a formáns fogalmát ettől a meghatározástól eltérően (is) értelmezzük, illetve figyelmet szentelünk a percepció szempontjából fontos (akusztikailag is mérhető) spektrális maximumhelyeknek, ezt megjegyzésként minden esetben feltüntetjük (vö. 2.1. fejezet).

Az alapprofundencia, a felharmonikusok, valamint a formánsok között fennálló összefüggés miatt kellően mély alapprofundenciák (például az átlagos felnött beszélők beszédhangjának) esetében a zöngére jellemző nagy felhangsűrűség miatt úgy tekinthetjük, hogy a formánsok mint a toldakécső rezonanciahelyei jó közelítéssel megállapíthatók a beszédhangok spektrumából, hiszen a mély alapprofundenciák esetében – amint láttuk – nagyobb számú felharmonikus esik egy-egy rezonanciahely sávzélességébe (tehát több felharmonikus vesz részt a spektrális csúszok kialakításában). Így alacsonyabb alapprofundenciákon az átviteli függvény jobb frekvenciafelbontású lesz a beszédhangok akusztikai lenyomatában (vö. WOLFE et al. 2009). Az alapprofundencia emelésével azonban, ami a jelen értekezés kérdésfelvetéseinek egyik alapvető problémakörét képezi, a felharmonikusok egyre növekvő távolsága miatt (vö. 6. és 7. ábra) a formánsok (mint rezonanciák) detektálhatósága, elérhetősége az akusztikai kimenetben mind az akusztikai mérések, mind pedig az emberi beszédpercepció számára egyre csökken.



7. ábra: A zöngés beszédhangok a forrás-szűrő modellben alacsony (A) és magas (B) alapprofundenciák esetén: fent a gégeben létrejött zöngé spektruma, középen a szűrőként funkcionáló átviteli függvény, alul pedig a szűrt, a szájnál kisugárzott akusztikai jel (EPPS et al. 1997: 1113 alapján)

A 7. ábra mély (A) és magas (B) alapprofrekvenciájú hangok vonalas spektrumát ábrázolja. Az ábrán jól látható, hogy míg alacsony alapprofrekvencián az F_1 és az F_2 rezonanciák egyaránt jól kivethetők az akusztikai kimenetben, magas alapprofrekvencián sem az F_1 , sem pedig az F_2 rezonancia frekvenciahelye nem detektálható nagy bizonyossággal. Az ábrán látható információk alapján mindösszesen annyit feltételezhetünk, hogy az f_0 , azaz a spektrum legalacsonyabb frekvencia-összetevője beleesik egy formáns sáv szélességébe, mert nagy intenzitást mutat.

A fentiek értelmében tehát azt mondhatjuk, hogy a toldalékcsőre jellemző sajátfrekvenciák vagy formánsok a magasabb alapprofrekvenciák esetében már nem feltétlenül egyeznek meg az akusztikai kimenetben megjelenő energiamaximumokkal – a kimenet egyes akusztikai összetevőinek intenzitása inkább csak következtetni enged arra, hogy az egyes felharmonikusok beleesnek-e az egyes formánshelyek sáv szélességébe. Így mind a formánsok mint rezonanciák frekvenciájára, mind pedig az azokat létrehozó artikulációs mozgásokra is egyre kisebb biztonsággal tudunk következtetni az alapprofrekvencia emelésével.

A fent leírt jelenséget, mely tehát a toldalékcső átviteli függvényének az alapprofrekvencia emelésével egyre alacsonyabb frekvenciafelbontását eredményezi, egyes kutatók a beszédhangok magas alapprofrekvencián tapasztalható **spektrális alul-mintavételezettségének** nevezik (DE CHEVEIGNÉ–KAWAHARA 1999). Bár ez a terminus megtévesztő lehet, ha a mintavételezés fogalmát mereven a digitális jelfeldolgozásban használatos értelemben próbáljuk alkalmazni (vö. pl. SZASZÁK 2010), mégis az analógia nagyon szemléletesen és egyszerűen ragadja meg a „mintavételezési frekvenciából” (azaz itt a felharmonikusok távolságából, illetve az alapprofrekvenciából) fakadó felbontásbeli különbségeket az alapprofrekvencia függvényében. Éppen ezért a magas alapprofrekvenciájú hangok spektrális alul-mintavételezettségére a jelen értekezésben is hivatkozni fogunk, a formáns fogalmát pedig a toldalékcső rezonanciáira alkalmazzuk.

1.1.1.3. A MAGÁNHANGZÓK ARTIKULÁCIÓS ÉS AKUSZTIKAI JELLEMZÉSE

A magánhangzókat artikulációsan és akusztikailag is jellemezhetjük: az előbbi a magánhangzók létrehozását előidéző artikulációs mozgásoknak, az artikulátorok állásának a megjelölése, az utóbbit pedig egyebek mellett – a fent említett csőmodell alapján – az artikulációval szorosan összefüggő formánsfrekvenciáknak a megadása jelenti. Az artikulációs és akusztikai meghatározás célja, hogy elkülönítsük és jellemezzük az egyes magánhangzóminőségeket. A **magánhangzó-minőség** terminus technicus az ejtett magánhangzó auditív minőségét jelöli – az a címke, melyet a három alapvető, a magánhangzók ejtését jellemző paraméter határoz meg: az elől vagy hátul képzettség, a nyitottság vagy függőleges nyelvhelyzet és az ajakműködés (ASHBY 2011). Éppen ezért ezek azok a paraméterek is, melyeket az artikulációs jellemzéskor megadunk. Az egyes paraméterek szorosan összefüggnek a toldalékcsőre a magánhangzók képzésekor jellemző sajátfrekvenciákkal, formánsokkal, aminek

eredményeképp a magánhangzók formánsai alkalmasak lesznek arra, hogy akusztikailag elkülönítsék és meghatározzák az egyes magánhangzó-minőségeket (vö. FANT 1960).

A magánhangzók képzésekor a hangszalagok (normál beszédmódban) mindig rezegnek,² az aktív és passzív artikulátorok, azaz a nyelv és a szájpad között létrejövő szűkület pedig megfelelően nagy ahhoz, hogy az artikulátorok ne érintkezzenek (tehát nem okoznak zárat, mint a zármássalhangzók, pl. a /t/ esetében), és hogy a szűkületen átáramló levegő ne okozzon turbulens zörejt (mint a résmássalhangzók, pl. a /s/ esetében). A zöngje jelenléte és az érintkezés vagy a zörej hiánya miatt a magánhangzókat a szonoráns hangok közé soroljuk (vö. pl. GÓSY 2004; ASHBY 2011).

Bizonyos értelemben a magánhangzók csoportosításának legalapvetőbb dimenziója a nyíltság. E szerint megkülönböztetünk magas(abb), illetve alacsony(abb) nyelvháttal képzett beszédhangokat (vö. STEVENS 1998). A nyíltságnak jellemzően négy fokozatát különítjük el, ezeket magyarul felső, középső, alsó és legalsó nyelvvállásnak nevezzük (GÓSY 2004). Amint azt a terminológia is mutatja, a nyíltság dimenzióját sokan elsősorban a nyelvtest függőleges helyzetének feleltetik meg. Valójában azonban a nyíltság kérdése az artikuláció, az akusztikum és a percepció együttes figyelembevételével ennél egy kicsit összetettebb jelenségnek látszik, melyben nem csak a nyelvtest, de az állkapocsnyitás szöge is fontos szereppel bírhat. Erről még a következő alfejezetben a magyar magánhangzók ismertetésekor valamivel bővebben is szólunk.

Az artikuláció akusztikai vetületét vizsgálva azt tapasztaljuk, hogy míg a magasabb nyelvállású magánhangzók F_1 -értéke alacsonyabb, az alacsonyabb nyelvállás következményeképp az F_1 értéke magasabb lesz. Ha a nyelvtest magasabban helyezkedik el, a garatüreg térfogata a neutrális nyelvhelyzethez képest nagyobb lesz, mely esetben az F_1 értékét elsősorban a létrejött szűkület mérete, hossza határozza meg. Az F_1 fiziológiailag meghatározott minimális értékének eléréséhez a nyelvtest emelése mellett az ajkak réses nyitása és az ajaknyílás keskenyítése, hosszítása szükséges. Mivel ez az artikulációs konfiguráció éppen az /i:/ hangra jellemző, ezt a hangzót az artikulációs és akusztikai magánhangzótér egyik sarokpontjának tekintjük (STEVENS 1998).

Az állkapocs nyitása és a nyelvhát süllyesztése emeli az F_1 értékét. A nyugalmi értékekhez képest az F_1 növekedését okozza ugyanis, ha a nyelv lefelé és kicsit hátrafelé, a garat felé húzódik, melynek következtében a szűkület a szájüreg hátsó részében, a garatnál jön létre. Ezt elsősorban az idézi elő, ha a nyelvtest elülső része lefelé mozdul, de a hatást erősíti, azaz az F_1 értékét tovább növeli, ha megnövekszik az szájüreg elülső részének térfogata is, melyet az állkapocsnyitás szögének növelésével idézhetünk elő (STEVENS 1998).

Az F_1 értéke tehát összefoglalóan a nyelvtest függőleges helyzetével, valamint az állkapocs nyitásszögével függ össze, így a magánhangzók nyíltságát artikulációsan a nyelvtest és/vagy az állkapocs nyitásszögével, akusztikailag pedig az F_1 értékével jellemezzük.

² Bár folyamatos beszédben a hangzók egymásra hatásának következtében előfordulhat a magánhangzók zöngétlenedése, a zöngétlen magánhangzókat a magyarban nem tekintjük önálló, fonémaszerepű hangoknak.

Az F_2 értékére elsősorban a nyelv vízszintes helyzete hat, azaz az a körülmény, hogy a nyelvtest elöl vagy hátul helyezkedik-e el a szájüregben. Általánosan a nyelvtest előretolulása az F_2 értékének növekedését eredményezi, míg a nyelvtest hátrafelé pozicionálása az F_2 értékének csökkenését okozza. Igaz azonban az is, hogy az F_2 maximuma a felső nyelvtámasztások esetében magasabb lehet, mint az alsó nyelvtámasztásoknál (STEVENS 1998). STEVENS (1998) modelljének adatai alapján megállapítható, hogy a felső(bb) nyelvtámasztás magánhangzók F_2 -értékének hatékony csökkentéséhez szükséges az ajakkerekítés is, így a legzártabb (legalacsonyabb F_1 -gyel képzett) és egyúttal ajakkerekítéssel képzett /u/ képezi az artikulációs és akusztikai magánhangzótér egy másik sarokpontját (minimális F_1 - és F_2 -értékkel, legzártabb és egyúttal leghátrébb képzett hangként). Az ajakkerekítés hatása az F_2 értékére azt a tendenciát is magyarázza, mely szerint „jelöltebbek” a hátul képzett réses hangzók, azaz gyakoribbak azok a nyelvek, melyekben csak kerekítéses hátul képzett magánhangzók fordulnak elő. A hátul képzett hangzók ajakkerekítéses ejtése ugyanis a fentiek alapján akusztikailag motiváltabbnak tűnik, hiszen ez teszi lehetővé a legnagyobb akusztikai kontrasztot az elöl képzett hangokhoz képest.

Noha az alsó nyelvtámasztású hangok mozgásterét az elől-hátul képzettség dimenziója mentén fizikailag kisebbnek tetszhet a felsőbb nyelvtámasztású hangokénál, mégis lehetséges az elől és hátul képzettség distinkciója, mégpedig úgy, hogy a létrejövő szűkület helye az F_1 értékét csak minimálisan befolyásolja. Mégis, a legalsó nyelvtámasztású elöl képzett hangzók maximális F_1 - és F_2 -értékét az ajakrész ejtés tovább fokozhatja, ugyanis az ajakrész ejtés valamelyest emeli az F_2 értékét (STEVENS 1998).

A magánhangzók jellemzésének második artikulációs paramétere tehát a vízszintes nyelv helyzet. Ez a paraméter alapvetően két vagy három fokozatú lehet, beszélhetünk ugyanis elől képzett, hátul képzett vagy középen képzett hangokról. A vízszintes nyelv helyzet elsődleges akusztikai vetületének pedig az F_2 -t tekintjük, ugyanis az F_2 értéke elsősorban ezzel az artikulációs dimenzióval mutat szoros összefüggést. Láttuk azonban azt is, hogy az F_2 értékére valamelyest az ajkak réses vagy kerekítéses formálása is hatással van, hiszen a réses ejtés némileg emeli, az ajakkerekítéses ejtés pedig csökkenti az F_2 értékét. Ennek következtében az ajkak működése a harmadik, a magánhangzók jellemzésében általánosan megadott artikulációs paraméter, melyhez akusztikusan szintén az F_2 értékét társíthatjuk.

A fentiek rávilágítanak arra a korábban már említett tényre, hogy miért bír a toldalékcso első két sajátrezonanciája, azaz az első két formáns, az F_1 és az F_2 kitüntetett jelentőséggel a beszéddel foglalkozó tudományokban. Mivel az F_1 szoros összefüggést mutat a magánhangzók nyíltságával, az F_2 pedig a magánhangzók elől vagy hátul képzettségével, az F_1 és F_2 értéke könnyen (sőt sokszor félrevezetően könnyen) interpretálhatónak látszik a magánhangzók artikulációs jellemzőire nézvést. Emellett pedig az artikuláció és akusztikum ezen egyszerűnek látszó összefüggése a magánhangzó-minőségek kétféle (ha szigorúan nem is ekvivalens, de) szorosan összekapcsolható jellemzését is lehetővé teszi (az artikulációs és az akusztikai vetületben).

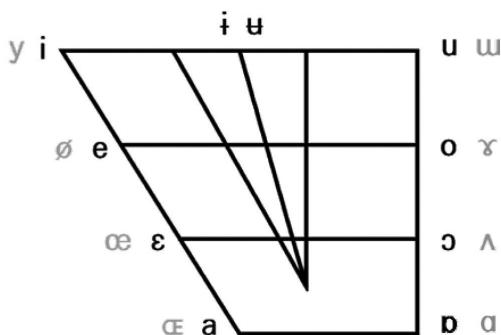
A teljesség kedvéért megemlítendő, hogy a fentiek mellett megkülönböztetünk még orális és nazoorális magánhangzókat (tehát a magánhangzók képzéséhez felvehetünk még egy artikulációs képzési jegyet), valamint monoftongus, diftongus és trifftongus magánhangzókat is. Ezek megkülönböztetése azonban a magyar köznyelvi magánhangzókészlet szempontjából másodlagos, hiszen a készlet minden eleme orális és monoftogus beszédhang. A nazoorális magánhangzók ejtése során (az orálisokkal szemben) a lágy száypad nyitva van (azaz nem zárja el a levegő útját az orrüreg felé), ezért az ejtés során a levegő nem csak a szájon, de az orron keresztül is távozik, ami sajátos szint ad a beszédhangoknak. A diftongusok (és trifftongusok) a monoftongusokkal szemben olyan magánhangzók, melyek képzési ideje alatt valamely, a magánhangzó minőségét meghatározó jegyben (pl. a nyelvallás fokában) egy (vagy a trifftongusok esetében egymás után több) változás történik. Bár a változás következtében valójában kettős- (vagy hármas-) hangzót ejtünk, de a két (vagy három) kiejtett hang nem képezhet például önálló szótagot, így nem is számít önálló fonémaértékkel bíró hangnak sem. A diftongusok és trifftongusok tehát nyelvileg egyetlen egységként, egyetlen fonémaként értelmezendők (GÓSY 2004).

Bár a fentiek ismeretében a magánhangzók rendszerezése ma már viszonylag egyszerűnek tűnhet, a rendszerezés szempontjai nem voltak mindig ilyen kézenfekvők. Amiatt ugyanis, hogy a magánhangzók képzésekor az artikulátorokat csak közelítjük, de nem érintjük egymáshoz, illetve a közelítés hatására nem jön létre zörej, valójában a magánhangzók jellemzése a mássalhangzóknál jóval nehezebb kérdés, két okból is. Egyfelől bár az előző alfejezetben említettük, hogy az egyes magánhangzók képzésekor a nyelv érintkezhet ugyan a nem mozgatható artikulátorokkal (pl. a nyelvperem a felső nyelvallású magánhangzók esetében hozzáér a felső fogsorhoz, vö. STEVENS 1998), mégis alapvetően beszéd közben nem kapunk jól megfigyelhető, pontos taktilis visszajelzéseket a magánhangzók létrehozását eredményező közelítés vagy szűkület létrejöttének helyéről, azaz a magánhangzó képzéshelyéről. Másfelől pedig a magánhangzók izolált ejtése statikus – szemben a mássalhangzókéval –, így kinesztetikus információk sem segítik a magánhangzók képzéshelyének azonosítását, miközben pedig tudjuk, a nyelv igen nagy variabilitással mozgatható és alakítható (ASHBY 2011).

Daniel JONES (1922) a magánhangzók jellemzésének problémáját megoldandó megalkotta a **kardinális magánhangzók** rendszerét, azaz viszonyítási pontokat határozott meg a magánhangzók képzéséhez használt artikulációs és akusztikai térben. JONES (1922) pusztán kivételesen jó hallására támaszkodva képes volt a jelen fejezetben korábban megfogalmazott artikulációs, akusztikai és a következőkben érintendő percepciós sajátosságok megragadására, és bizonyos szélső értékeket képező hangzóminőségek lehorgonyzásával rendszerbe szedte az emberi beszédképző szervekkel létrehozható magánhangzókat. A kardinális magánhangzók rendszere nem egy létező nyelv magánhangzóinak rendszere, hanem az artikulációs „munka-terület”, vagyis az **artikulációs magánhangzótér** egymástól – és ez egy igen fontos kitétel – auditve egyenlő távolságokra lévő kategóriákra osztott absztrakt struktúrája. A rendszerben megadott magánhangzók tehát maguk nem realizztikusak, céljuk csupán az, hogy olyan sarokpontokat képezzenek, melyekhez képest az egyes nyelvek realizztikus beszédhangjai

„lehorgonyozhatók” lehetnek. Ily módon a kardinális magánhangzók rendszere jó (auditív) mérőeszköze az egyes realiztikus magánhangzóknek: lehetővé teszi, hogy a beszédhangokat rendszerbe szedjük, és a különböző nyelvek magánhangzóit egymással összehasonlítható módon jellemezzük (vö. pl. ASHBY 2011).

Az első rendszer nyolc magánhangzó-minőséget tartalmazott, melyet aztán JONES megduplázott az ajakműködés szerint – minden kerekítéses hang mellé egy réses ejtésű párt rendelt és fordítva (lásd 8. ábra). Érdekes módon bár a kardinális magánhangzók rendszerében a magánhangzók elkülönítése hallásalapú, mégis a magánhangzókat leíró terminológia alapvetően artikulációs tulajdonságokon és terminológián alapszik (azaz a magánhangzókat artikulációsan jellemezzük és különítjük el egymástól). A rendszert többek közt éppen emiatt számos kritika érte. Mivel azonban az artikuláció (a függőleges és vízszintes nyelvhelyzet) és annak auditív eredménye, azaz a magánhangzó-minőség között erős korreláció tapasztalható, ezért ez a fajta megközelítés – normál beszédejtésben legalábbis – jó közelítésnek bizonyult a magánhangzók jellemzésére. Jól mutatja ezt az a tény is, hogy a magánhangzókat máig e szerint a rendszer szerint adjuk meg. Az első bővítést később továbbiak is követték, így alakult ki a jelenleg is használatos trapéz alakban vizualizált szisztéma.



8. ábra: Az elsődleges (fekete) és másodlagos (szürke) kardinális magánhangzók rendszere
(Forrás: <http://matters-phonetic.blogspot.hu/2011/02/romantic-cardinal-vowels.html>)

A kardinális magánhangzóknek a 8. ábrán bemutatott rendszere, azaz a magánhangzótrapéz a következőképpen olvasandó. A függőleges tengelyről a nyelvállás fokát vagy az állkapocsnyitás szögét olvashatjuk le: fent található a felső nyelvállású, zárt magánhangzók, alul pedig az alsó nyelvállású, nyílt magánhangzók. A vízszintes tengely a nyelv vízszintes pozícióját mutatja balra a palatális, azaz elől képzett, jobbra pedig a veláris vagy hátul képzett hangokkal. A vízszintes tengely mentén középtájon elkülönül a centrális vagy mediális képzéshely is. A magánhangzópárok bal oldali tagja minden esetben ajakkerekítéssel képzett, azaz labiális hang, míg a jobb oldali tag ajakréssel képzett, azaz illabiális.

Mivel a vízszintes és függőleges tengelyek az F_1 és F_2 értékével szoros összefüggésbe hozható artikulációs paramétereket ábrázolják, ezért a JONES (1922) rendszeréből kiindult, a nyelv-

helyzetet is figyelembe vevő, és ilyen értelemben inkább az artikulációs magánhangzóteret leképező IPA szabvány szerinti ábrázolás (vö. IPA Handbook, 1999) megközelítően jól reprodukálható az F_1 és az F_2 tengelyek mentén felvett **akusztikai magánhangzótérrel**. Bár ez utóbbi, tehát az akusztikai magánhangzótér nem feleltethető meg maradéktalanul az előbbinek, azaz az artikulációs magánhangzótérnek (például azért sem, mert az ajakkerekítés miatt alacsonyabb F_2 az akusztikai magánhangzótérben így elkerülhetetlenül a hátulképzettség „számlájára íródik”), a kettő elkülönítésének figyelembevételével az akusztikai magánhangzótér ábrázolása és vizsgálata jó és gyakran használt eszköz a magánhangzók megvalósulásának elemzésében.

A fent bemutatottak alapján tehát láthatjuk, hogy az egyes magánhangzó-minőségek artikulációs és akusztikai tekintetben elkülönülni látszanak egymástól, hiszen mind a nyíltság, mind a vízszintes nyelvhelyzet más-más akusztikai szerkezetet, eltérő F_1 - és F_2 -értékeket eredményez. Ebből kiindulva hagyományosan azt szokás feltételezni, hogy az egyes magánhangzók felismerése is alapvetően ezeken a kitüntetett szerepű frekvencia-összetevőkön múlik. A valóságban azonban az F_1 - és F_2 -értékek igen nagy variabilitást mutathatnak, és sokszor átfedésben is vannak az egyes magánhangzó-minőségek között (legalábbis a lineáris fizikai mértékegységen, azaz Herten alapuló ábrázolásban). A magánhangzók észlelése vagy percepciója, azaz az egyes magánhangzó-minőségek azonosítása tehát korántsem triviális kognitív működések eredménye. A magánhangzók észlelésével kapcsolatos ismereteinkre a következő, a magyar magánhangzók bemutatását célzó alfejezet után térünk ki.

1.1.1.4. A MAGYAR KÖZNYELVI MAGÁNhangzók ARTIKULÁCIÓS ÉS AKUSZTIKAI JELLEMZÉSE

A magyar magánhangzók (artikulációs) jellemzéséhez hagyományosan (valamint az előző alfejezetben bemutatottaknak is megfelelően) a nyelv nyugalmi pozíciójához viszonyított vízszintes és függőleges helyzetét adjuk meg, valamint az ajakaktivitást jellemezzük. Ezen felül pedig, mivel a magyarban jelentésmegkülönböztető szerepe van, szokás megadni a magánhangzók nyelvi, azaz fonológiai hosszúságát is (vö. GÓSY 2004).

A nyelv függőleges helyzete szerint megkülönböztetünk felső, középső, alsó és legalsó nyelvállásfokokat, a nyelv vízszintes helyzete szerint elől és hátul képzett vagy másként palatális és veláris hangzókat, illetve középen képzett, azaz centrális vagy mediális hangzókat, az ajakműködés szerint elkülönítünk ajakréssel képzett réses vagy illabiális és ajakkerekítéssel képzett kerekített vagy labiális hangzókat, míg a nyelvi hosszúság tekintetében megkülönböztetjük a fonológiaiilag hosszú és rövid hangokat (lásd BOLLA 1995; KASSAI 1998; GÓSY 2004; MÁDY 2008).

A magyar köznyelvi magánhangzóknek a nyelvállásfok szerinti besorolási lehetőségeit az *1. táblázatban* foglaljuk össze. A táblázat csak a minden idézett leírásban egyaránt meglévő 14 magánhangzót tartalmazza. Mivel ezek azok a magyar magánhangzók, melyek a leírások

szerint nem kötődnek bizonyos beszédhelyzetekhez vagy nyelvváltozatokhoz, így ezek között nem szerepel a BOLLA (1995) és KASSAI (1998) által feltételezett zárt *ɛ*, azaz a rövid /e/, valamint a GÓSY (2004) által feltételezett rövid /a/ és a svá (/ə/).

1. táblázat: A magyar köznyelvi (nyelvváltozatokhoz vagy beszédmódokhoz nem kötődő megjelenésű) magánhangzók nyelvállásfok/nyíltság szerinti fonetikai* besorolása a különböző szakirodalmi forrásokban (az első oszlopban a nyíltság mértékének az értekezésben is több helyütt hivatkozott számszerű kifejezését adtuk meg)

Nyíltsági fok (TRAUNMÜLLER 1981)	A nyelvállás foka	BOLLA (1995)	KASSAI (1998)	SZENDE – IPA Handbook (1999)	GÓSY (2004)	MÁDY (2008)
1.	Felső	i i: yy: u u:	i i: y y: u u:	i i: y y: u u:	i i: y y: u u:	i i: y y: u u:
2.	Középső	e: ø ø: o o:	e: ø ø: o o:	e: ø ø: o o:	e: ø ø: o o:	e: ø ø: o o:
3.	Alsó	ɛ ɔ	ɛ ɔ	ɛ	ɛ ɔ	ɛ
4.	Legalsó	a:	a:	a: a/ɒ	a:	a: ɒ

* A fonológiai besorolás talán legnagyobb eltérése a fonetikai besoroláshoz képest a nyelvállásfok tekintetében, hogy nem különböztet meg alsó és legalsó nyelvállásfokokat, az /e a: ɔ/ magánhangzókat egyszerűen alsónak ([–high][+low] jegyűnek) tekinti (vö. SIPTÁR–TÖRKENCZY 2000).

Noha az itt idézett szövegekben a magánhangzók nyíltsági korrelátumának alapvetően a nyelvállás fokát tekintik a szerzők (így ezt szerepeltettük a táblázat második oszlopában is), egyes leírások emellett (gyakorlatilag a nyelvállásfok szinonimájaként) megadják 1. a szájüreg nyíltsági fokát is az állkapocs nyitásszögében kifejezve, azaz (a nyelvállásfokok fenti felsorolásának megfelelően) rendre a zárt, félig zárt, félig nyílt és nyílt állkapocsnyitást, valamint 2. az ajkak által képzett nyílás tágasságát is, mely értékei rendre: legszűkebb, szűk, tág, legtágabb (lásd pl. KASSAI 1998). A nemzetközi szakirodalomban ez a gyakorlat egyébként egyáltalán nem szokatlan, amint azt később látni fogjuk, WOOD (1975) és TRAUNMÜLLER (1981) is az állkapocsnyitás szögét jelöli meg mint a magánhangzó nyíltságának elsődleges artikulációs korrelátuma (WOOD (1975) szerint ehhez a nyelvállásfok mint a feszes-laza szembenállás artikulációs korrelátuma járul). TRAUNMÜLLER (1981) ráadásul a magánhangzók percepciósi nyíltsága kapcsán is a nyíltság, nyitásfok vagy nyíltsági fok kifejezéseknek fordítható *openness* és *degree of openness* terminusokat használja.

TRAUNMÜLLER (1981) érvelése szerint a magánhangzók a nyíltság szerint észlelt distinktív jegyeinek (azaz a különböző nyíltsági fokoknak) a megragadása az egyszerűség és egyértelműség kedvéért szükségszerűen artikulációs terminusokkal (azaz a nyelvállás, az állkapocsnyitás vagy az ajaknyitás megadásával) kell, hogy éljen. Mindemellett azonban a legjobb fonológiai általánosíthatóságot, véli, az állkapocsnyitásszög megjelölése adja, hiszen a nyelvállás valójában a veláris és palatális, azaz az elől és hátul képzett hangok esetében némileg

eltér, azaz variábilis (vö. LADEFOGED 1967; idézi TRAUNMÜLLER 1981; GÓSY 2004). A jelen értekezésben, hogy elkerüljük az artikulációs, akusztikai vagy percepciós nyíltság megragadásának lehetséges dilemmáit, a magyar fonetikai hagyományban megszokott nyelvvállásfok terminus mellett alternatívaként hivatkozni fogunk a magánhangzók „nyíltságára”, valamint a nyíltság számszerűen kifejezett mértékére is (TRAUNMÜLLER 1981 nyomán). A nyíltsági fokok számszerűsített értékeit az 1. táblázat első oszlopában tüntettük fel.

Az idézett források közti fő különbséget az *a* graféma által jelölt hang átírása és besorolása jelenti, azaz a szerzők nem egységesen feleltetik meg azt egy adott kardinális magánhangzónak: az *a*-t az *á*-val (azaz /a:/-val) mint legelső nyelvvállású hanggal szembeállítva több leírásban is alsó nyelvvállású hangként kategorizálják, míg más források legelső nyelvvállásúként tartják számon. A lehetséges megjelölések szerint tehát az *a* graféma által jelölt hang vagy ajakkerekítéssel képzett alsó nyelvvállású /ɔ/, vagy ajakkerekítéssel képzett legelső nyelvvállású /ɒ/, vagy ajakréssel képzett legelső nyelvvállású /a/ magánhangzó. A hangzó vitatott besorolásával kapcsolatosan legújabban MÁDY (2008) végzett artikulációs vizsgálatot elektromágneses artikulográf segítségével. Eredményei szerint – bár az állkapocsnyitás kis mértékben valóban nagyobb volt az /a:/ esetében, mint a kérdéses hangzónál – a két magánhangzó esetében megfigyelt nyelvvállásfok gyakorlatilag megegyezett, miközben az *a* graféma jelölte magánhangzó esetében intenzív ajakkerekítés volt megfigyelhető. MÁDY következtetése szerint ezek az eredmények a többi magánhangzó ejtési mintázataihoz is viszonyítva, valamint korábbi vizsgálatok tanulságai alapján az *a* hang legelső, az /a:/-val egyező nyelvvállásfokú képzésére utalnak. A szerző érvelésében WOOD (1975) munkáját idézi, akinek megállapítása szerint az ajakkerekítéses hangzók a szűkebb ajaknyílás következtében természetesen kisebb állkapocsnyítással is képzettek az azonos nyelvvállásfokú párjaiknál, ennek megfelelően pusztán az állkapocsnyításban tapasztalt különbségek nem jelentik, hogy az adott magánhangzókat nyíltság vagy nyelvvállásfok tekintetében különbözőnek kellene tekinteni. A jelen értekezésben ezen eredmények figyelembevételével az *a* grafémával jelölt beszédhangot legelső nyelvvállásúként soroljuk be és ennek megfelelően /ɒ/-ként jelöljük (tehát a SZENDE (1999) és MÁDY (2008) által javasolt besorolást alkalmazzuk).

A magyar köznyelvi magánhangzók leírása a nyelv vízszintes helyzetének szempontjából nagyjából egységesnek mondható a szakirodalomban. A fent idézett források a veláris hangzók közé sorolják az /u u: o o: ɒ a:/ magánhangzókat, míg a palatálisok közé az /i i: y y: e: ε ø ø:/ magánhangzókat. A rendszer egyetlen valamelyest bizonytalannak tekinthető pontját az /a:/ és annak státusza képezi, mert bár az /a:/ akusztikailag leginkább centrálisnak vagy mediálisnak tekinthető – erre reflektál például BOLLA (1995), SZENDE (1999) vagy GÓSY (2012) is – fonológiai viselkedése miatt hagyományosan a velárisokhoz sorolják. (Az /a:/ veláris hangként viselkedik a magánhangzó-harmóniában, valamint a kétségtelenül veláris /ɒ/-val váltokozva vesz részt a fonológiai alternációkban is, vö. pl. SIPTÁR–TÖRKENCZY 2000; SIPTÁR 2014).

Az ajakműködés tekintetében viszonylag nagy az egyezés BOLLA (1995), KASSAI (1998), SZENDE (1999) és GÓSY (2004) között. Bár a szerzők mindegyike reflektál rá, hogy az alaptípusoknak számító ajakrészes és ajakkerekítéses ejtés nyelvvállásfokoként az állkapocsnyítás

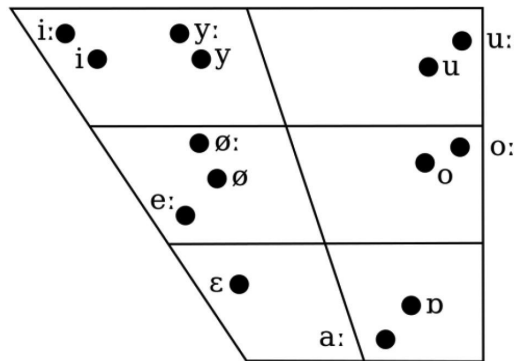
mértékétől függően változó megjelenést, változó ajakformát, elérő mértékű ajaknyitást mutat, a magánhangzókat egyértelműen besorolják az ajakkerekítéses és ajakréses csoportokba. A felosztás a következőképpen alakul. Ajakkerekítéses vagy labiális hangzók: /u u: o o: ɒ y y: ø ø:/, valamint ajakréses vagy illabiális hangzók: /i i: e: ε a:/. Ahogyan arra BOLLA (1995) is utal, az alsó vagy legalsó nyelvállásfokon ejtett labiális és illabiális magánhangzók közti különbség kisebbnek vagy legalábbis erősen más jellegűnek mutakozhat, mint a felsőbb nyelvállások esetében. Valószínűleg ez eredményezi azt is, hogy például SZENDE (1999) a vitatott minőségű *a* graféma által jelölt hangot elsődlegesen nem ajakkerekítéses /ɒ/, hanem ajakréses /a/ hangként írja át, majd megjegyzi, hogy mivel a hangzó valójában mutat némi kerekítést, így akár az /ɒ/-ként való átírása is lehetséges lenne. A jelen értekezésben, amint azt már fentebb is említettük, az *a* graféma által jelölt hang megjelölésére az ajakkerekítéses /ɒ/-t használjuk.

A magánhangzó-minőséget alapvetően meghatározó hangszínen (azaz az eddig taglalt három paraméteren) felül a magánhangzók jellemezhetőek a nyelvi, azaz fonológiai hosszúságukkal is. Különösképp fontos jegy ez a magyarban, ahol a fonológiai hosszúság jelentésmegkülönböztető szereppel bír (vö. pl. SIPTÁR–TÖRKENCZY 2000). Ennek megfelelően a magánhangzók (az *I. táblázatban* is szemléltetett módon) a fonológiai hosszúság szerint párokba rendezhetőek. Amennyiben csak a fent idézett leírásokban egyaránt megjelölt (azaz a táblázatban feltüntetett), köznyelvi magánhangzókat vizsgáljuk, a 14 magánhangzó 7, fonológiailag a hosszúság szerint opponáló párba állítható. Bár az /e:/ és /ε/, valamint az /ɒ/ és /a:/ a hangszínekben is jelentősen eltérnek egymástól, mégis fonológiai szempontból pároknak tekintethetők. Ezt a megfontolást elsősorban az indokolja, hogy e hosszúsági párok tagjai nyelvi viselkedésüket tekintve a csak a nyelvi időtartamukban eltérő párokhoz (pl. az /i i:/-hez) hasonlóan szisztematikusan (regulárisan) váltakoznak, azaz alternálnak (pl. a *víz* – *vizek*, *kéz* – *kezek*, *nyár* – *nyarak* alakokban, vö. SIPTÁR–TÖRKENCZY 2000). Ugyanakkor éppen a hangszínbeli különbséget szem előtt tartva az utóbb említett beszédhangokat rövid vagy hosszú pár nélküli hangzóknak tekinti például GÓSY (2004). Az ő fonetikai szempontú rendszerezésében ugyanis csak a hangszínbeli különbséget nem mutató beszédhangok állíthatók egymással hosszú-rövid párokba.

A jelen értekezés szempontjából kiemelendő, hogy a hangszínbeli és időtartambeli különbségeket is figyelembe véve úgy tekinthetjük, hogy a magyarban mindösszesen 9-féle magánhangzó-minőség különül el, melyek (a fonológiai párok hosszú tagjával szemléltetve) a következők: /i: y: u: e: ø: o: ε a: ɒ/. Az ebből a felsorolásból kimaradt felső és középső nyelvállású rövid beszédhangok minősége (elméletileg) megegyezik a megfelelő hosszú pár minőségével, és a rövid-hosszú párok közti különbséget – ismét elsősorban elméletileg – csak a nyelvi hosszúság (vagy másként a nyelvi kvantitás) adja. Valójában természetesen a helyzet nem ennyire egyszerű, ugyanis a hosszú magánhangzók hangszínekben (a nyelvállás fokától is erősen függő módon) többé vagy kevésbé minden esetben eltérnek rövid párjaiktól (vö. KASSAI 1998; GÓSY 2004). Ez a jelenség (és az abban rejlő fonológiailag is releváns probléma) jól megragadható a H&H elmélet keretei között (az elmélettel kapcsolatosan lásd a jelen kötet *1.1.2.1.*

alfejezetét). A H&H elmélet értelmében a rövid-hosszú párok esetében tapasztalt hangszínbeli különbségek elvben a következő két ok valamelyikére vezethetők vissza (vö. MÁDY 2008). 1. A magánhangzópárok célkonfigurációja megegyezik, a hangzók közti hangszínbeli különbség pusztán az alulkonfiguráltság vagy más szóval a célalulmúlás eredménye, hiszen a rövid magánhangzók képzésére fordított rövidebb idő miatt a beszélő nem képes elérni a hosszú magánhangzókra jellemző konfigurációt. Ez a feltételezés akkor valószínűsíthető, ha a rövid hangzók centralizáltabbak hosszú párjuknál, azaz a hosszúaknál valamivel jobban a magánhangzótér közepe felé tolódnak. Ez a középső és felső nyelvvállású rövid magánhangzók akusztikai adatai alapján általában sejthető tendencia (vö. 9. ábra). 2. A rövid-hosszú magánhangzópárok között tapasztalt hangszínbeli különbség nem célalulmúlás eredménye, hanem annak a lenyomata, hogy a pár tagjainak célkonfigurációja eredendően eltér. Bár ezt eddig artikulációs vizsgálatokkal még nem igazolták, a magyar esetében feltehető, hogy a felső és középső nyelvvállású rövid magánhangzók artikulációs célkonfigurációja megegyezik a hosszú párjukéval (azaz a középső és felső nyelvvállású rövid és hosszú magánhangzók közti hangszínbeli különbség mindösszesen az artikulációra fordított idő következménye, vö. MÁDY 2008). Éppen ezért a jelen értekezés erejéig magunk is elfogadottnak tekintjük azt a feltételezést, hogy a középső és felső nyelvvállású magánhangzók rövid párjai csak az alulkonfiguráltság miatt különböznek hangszínbéli a megfelelő hosszú pártól, ezáltal azt is feltételezzük, hogy a magyarban 9-féle magánhangzó-minőséget különböztethetünk meg.

Az értekezésben feltételezett magánhangzókat a 9. ábra összesíti az IPA nemzetközi szabványának (IPA Handbook, 1999) megfelelő, trapézios elrendezésben, mely az absztraktabb jellemzőket, valamint az artikulációs és az akusztikai sajátságokat együttesen ábrázolja.



9. ábra: A magyar magánhangzóknak az értekezésben feltételezett rendszere a hagyományos, trapéz formájú elrendezésben (Forrás: Wikipedia Commons, vö. SZENDE 1994: 92)

A magánhangzótrapéz horizontális dimenziója az előző alfejezetben már bemutatott módon a nyelv vízszintes irányú helyzetével függ össze: bal oldalon az elől képzett, míg jobb oldalon a hátul képzett hangokat találjuk (az oldalakat a centrális képzéshelyet jelölő függőleges vonal

osztja ketté). A függőleges irányú dimenzió pedig, ismét csak a JONES (1922) rendszere kapcsán már taglaltaknak megfelelően, a nyelv függőleges helyzetére utal: legfelül a felső nyelvállású/zárt/első nyíltsági fokú magánhangzókat, legalul pedig a legalsó nyelvállású/legnyíltabb/negyedik nyíltsági fokú magánhangzókat találjuk. Egy adott oldalon (tehát a palatális vagy veláris csoportokon belül) és egy adott nyelvállásfokon belül bal kéz felé találjuk az ajakréses, míg jobb kéz felé az ajakkerekítéses hangzókat.

1.1.2. A magánhangzók észlelése

A pszicholingvisztikai modellekben a nyelvi feldolgozás legkisebb egységeként általában a hangszintű vagy más szóval szegmentális szintet említik. A legtöbb modellben ehhez a szinthez tartozik egy absztraktabb (fonológiai) és egy, a fizikai valósághoz közelebb álló (fonetikai) szint. Az ezekhez a szintekhez tartozó elemzési folyamatokat nevezzük átfogóan **beszédészlelési** vagy **beszédpercepciós** folyamatoknak – míg a beszéd magasabb szintű egységeinek megértését, a szavak, mondatok, megnyilatkozások feldolgozását is felölelő folyamatokat átfogóan a **beszédmegértés** folyamatának nevezzük (vö. GÓSY 2004).

Bár a nyelvi absztrakció, azaz a fonológia szintjén csak igen korlátozott számú hangot feltételezünk (azaz úgy véljük, hogy az egyes nyelvek csak viszonylag korlátozott számú fonémával gazdálkodnak, vö. pl. a jelenleg is érvényben lévő nemzetközi fonetikai ábécét, az IPA-t, mely mindösszesen huszonnyolc magánhangzó-fonémát különböztet meg, vö. IPA 2015), a beszédhangok a fizikai valóság, azaz a fonetika szintjén igen változatos megjelenésűek, azaz variábilisak lehetnek. A variabilitás fakadhat például a beszélő életkorából (a toldalékcső méretének sajátosságaiból), a beszédhang hangkörnyezetéből, illetve abból is, hogy gyors vagy lassú beszédben, hangsúlyos vagy hangsúlytalan szótagban valósul-e meg az adott beszédhang.

A beszédhangok észlelése, percepciója, azonosítása azt jelenti, hogy az emberi beszédfeldolgozás valamilyen módon a számtalan lehetséges hangzómegevalósulás egy-egy példányát a viszonylag kisszámú absztrakt és invariáns egységekhez, a fonémákhoz társítja. A beszédészlelési modellek egy része azt feltételezi, hogy a beszédészlelési mechanizmus ehhez a társítási processzushoz kiszűri a változatos megevalósulások szisztematikus, de fonetikailag-fonológiai nem distinktív jegyeit (amelyek felismerése, úgymond, nem segítené hozzá a hallgatót a beszédhang beazonosításához), és kiemeli, felismeri az invariáns jegyeket, melyek a fonetikai-fonológiai kategóriákat meghatározzák. Ezt a folyamatot az itt megfogalmazott meghatározásból következő módon a szakirodalom gyakran normalizálásnak is nevezi (NEARY 1989). NEARY (1989) „normalizálandó” hatásokként elsősorban a beszélő-, valamint a kontextusfüggő hatásokat említi, tehát megkülönbözteti a beszédhangoknak a beszélő egyéni ejtési és fiziológiai adottságaiból adódó, valamint a beszédhangnak a hangkörnyezetből fakadó (például a szomszédos mássalhangzó képzéshelye miatt jellemző) variabilitását. A beszédpercepciós elméletek (és vizsgálatok) jó részének fókuszában az itt

említett normalizási folyamat áll, valamint az a kérdés, hogy melyek azok az invariáns jegyek, amelyek felismerési kulcsokként szolgálnak egy-egy beszédhang (absztrakt) kategóriába sorolásához.

A beszéd feldolgozása, ahogyan azt már említettük, nem áll meg a hangszintű elemzésnél – a beszédmegértés teljes folyamatának és az ahhoz kapcsolódó elméleteknek az ismertetésére azonban a jelen kötet keretei között nincs módunk. Mivel a jelen vizsgálatokban elsősorban csak a szegmentális szintű beszédhang-azonosítással foglalkozunk, a jelen fejezet elméleti bevezetője is csak ennek kérdéseire fókuszál.

Mielőtt ismertetnénk a magánhangzók percepciójával kapcsolatos fontosabb kutatási eredményeket, röviden bemutatunk néhány, a beszédhangok percepciójával kapcsolatos alapvető elméletet. Tesszük ezt egyrészt azért, hogy megteremtsük a magánhangzókra vonatkozó eredmények értelmezéséhez szükséges elméleti keretet, másrészt pedig azért is, hogy szemléltessük a beszédhang-azonosítással kapcsolatos eddigi ismereteink sokféleségét – és bizonytalanságait is.

1.1.2.1. BESZÉDPERCEPCIÓS ELMÉLETEK

A beszédpercepció modelleket első közelítésben három nagy csoportba oszthatjuk – a három csoport aszerint oszlik meg, hogy az egyes percepció elméletek miben vélik biztosítottnak a beszédészleléshez szükségesnek tartott invarianciát. Az első a) csoportba tartozó modellek az invarianciát a magánhangzók artikulációjában, a második b) csoport elméletei a magánhangzók akusztikumában, míg a harmadik c) csoportba tartozó elméletek az észlelőben, az észlelő percepció mechanizmusaiban rejlenek feltételezik³ (vö. LINDBLOM et al. 1992).

a) Az első csoportba tartozó egyik legismertebb elmélet a **beszédészlelés motoros elmélete** (motor theory of speech perception). A modell elsősorban LIBERMAN nevéhez és munkásságához (valamint a Haskins Laboratories intézményhez) kötődik, az elmélet kifejtését (pontosabban annak javított változatát) LIBERMAN és MATTINGLY publikálta 1985-ben. A motoros elmélet szerint a beszéd észleléséhez a kogníció egy szerkezetileg különálló egységet, modult tart fenn, melynek egyik feladata, hogy a beszélők által ejtett variábilis magánhangzó-megvalósulásokból dekódolja a hallgató számára a beszédhang létrehozásakor a beszélő toldalékcsövére jellemző artikulációs konfigurációt, illetve artikulációs gesztusokat. A motoros elmélet azt feltételezi, hogy a beszéd során létrejövő akusztikai jel transzformációján keresztül a hallgató valójában az artikulációs információt detektálja, és ennek megfelelően az agy sem a hangok akusztikai lenyomatát, hanem a beszédhangok ejtéséhez szükséges motoros parancsokat (artikulációs gesztusokat) reprezentálja (mégpedig invariáns módon). A tükörneuronok

³ LINDBLOM (1990) egyébként az első két elméletre átfogóan jelalapúakként (*signal-based*) hivatkozik, míg az észlelő percepció mechanizmusaiban rejülő invarianciát feltételező elméletet, azaz saját H&H elméletét a jelfeldolgozást kiegészítő vagy jelfeldolgozáson túli folyamatokon (*signal-complementary processes*) alapulónak tekinti.

felfedezésével a kilencvenes években ez az elmélet új erőre kapott. A tükörneuronok olyan agyi idegsejtek, melyek mind egy adott cselekvés végrehajtása (pl. egy tárgy megragadása), mind pedig a cselekvés megfigyelése (a tárgy megragadásának látványa) közben „tüzelnek”.⁴ Noha ez a megfigyelés önmagában csak annyit jelent, hogy a szenzoros és motoros működések összefüggése egyfajta korrelációként jellemezhető (mert a motoros területek aktiválódnak a percepciós folyamat közben), mégis ezt az összefüggést a szakirodalom gyakorta kauzalitásként értelmezi, és azt állítja, hogy a motoros aktiváció célja az észlelési mechanizmus segítése (vö. HICKOK 2010). Bár a felfedezés erős empirikus bizonyítéknak tűnik a motoros teória helytállóságára, a szenzoros, azaz percepciós működések közbeni motoros aktivációt, azaz a tükörneuronok működését egyes vizsgálatok a fentiekől eltérően magyarázzák. HICKOK (2010) szerint a tükörneuronok működése elsősorban csak a szenzoros és motoros területek asszociációját mutatja, nem pedig azt, hogy a motoros aktiváció a percepció szükséges része lenne (hiszen a korreláció nem értelmezhető törvényszerűen kauzalitási, oksági viszonyként). A HICKOK által javasolt értelmezés szerint a percepció közbeni motoros aktiváció valójában az észlelés közben meginduló előkészítő folyamatok korrelátuma, azaz annak a lenyomata, hogy a megfigyelő a megfigyelt cselekvésre válaszreakciót (például a beszéd esetében válaszmegszólalást) tervez. Bár a motoros teória állításai sok szempontból vitathatók, és az elméletet sok kritika éri, mégis számos jelenség (pl. a koartikuláció) megragadására sok szempontból igen alkalmasnak bizonyul, így a mai napig is használt, népszerű beszédpercepciós modellek egyikeként tarthatjuk számon.

b) Az akusztikai (és bizonyos értelemben a percepciós) invarianciára alapozó elméletek legismertebbje az elsősorban STEVENS nevéhez (és a Massachusetts Institute of Technology, MIT intézményhez) kötődő **kvantális elmélet**. STEVENS (1989) szerint az artikuláció és az akusztikum közti összefüggés nem lineáris, hanem ugrásszerű, azaz kvantális, tehát az akusztikai jel bizonyos mértékű és jellegű artikulációs eltérésekre érzéketlenebb, míg mások nagyobb változásokat okoznak benne. A kvantális elmélet tehát kategóriákat, valamint kategoriális észlelést feltételez, ami – ilyen módon – megkönnyíti az észlelés folyamatát. Az elmélet számos további beszédpercepciós elképzeléssel, kísérleti eredménnyel kapcsolatba hozható (lásd pl. CHISTOVICH és LUBLINSKAJA (1979) eredményeit, melyekről részletesebben a következő alfejezetben lesz szó), valamint abból számos további percepciós teória is kinőtt. Egy ilyen elmélet például a LAFF (Lecixal Access from Features, 'lexikai hozzáférés jegyek alapján'), mely szerint a hallgató az észlelés során nem fonémákat azonosít, hanem (bináris) distinktív (viszonylag egyszerű akusztikai korrelátumokkal rendelkező) jegyeket különböztet meg, és a beszédjelet, a szavakat e jegyek sorozataként észleli. Jegyen az elmélet képviselői nem a beszédjel hullámformájában egyértelműen elhatárolható szakaszokat vagy paramétereket értenek, hanem absztrakt egységeket: egy példa a [+high] azaz [+felső] jegy, amely a függőleges nyelv helyzetre utal. Az egyes szavak felismerését, azaz aktivációjukat a LAFF elmélet szerint a mentális lexikonban az facilitálja, hogy a detektált jegysorozatot

⁴ Azaz aktivációt mutatnak, reagálnak.

az észlelő folyamatosan összeveti a mentális lexikonban tárolt mintázatokkal, majd a legkisebb különbség elvén „dönt”, és a szó aktiválódik (STEVENS 1989, 2002). A LAFF előnye, hogy mivel az észlelést a fonémák azonosítása helyett a jegysorozatok azonosításában fogalmazza meg, tudja kezelni azt is, hogy az egyes jegyek akár időbeni átfedéssel is megjelenhetnek, tehát az elmélet alkalmas a koartikulációs folyamatok megragadására is. Ráadásul a modell kísérletes úton jól tesztelhető – bár az is igaz, hogy a modell állításaival kapcsolatos empirikus vizsgálatok szinte kizárólag csak mesterségesen (azaz laborban) előhívott beszédre alapulnak, így kérdéses, hogy vajon a spontán beszédben is érvényesek-e (MÁDY 2013).

c) A harmadik, az invarianciát a hallgatóban feltételező irányzat híres képviselője Björn LINDBLOM, aki kutatási eredményeire támaszkodva megfogalmazta a **H&H elméletet**. LINDBLOM (1990) úgy véli, az állandóság nem az akusztikai vagy artikulációs jelben, hanem a hallgatóban keresendő, ugyanis az állandóság nem akusztikai vagy artikulációs jellemző, hanem a beszéd reprezentációjának tulajdonsága, a hallgató pedig nem „üres doboz”, ugyanis számos, az aktuális beszédjelen kívüli információ is a rendelkezésére áll. A H&H név a *hyper-* és *hypospeech* kifejezések rövidítéséből áll elő, melyek a fonetikai információban gazdag és az információhiányos jel koncepcióit fedik (egy kontinuum két végpontjaként). LINDBLOM elmélete szerint a beszélő a beszéd interakciós helyzetében az információgazdagság és információhiány skálája mentén rendelkezésére álló lehetséges megvalósítási formák közül választ annak fényében, hogy mit feltételez a hallgató számára (a jeltől függetlenül) elérhető információról. LINDBLOM szerint a beszéd leképezésére szolgáló komplex reprezentációinkat azon jelekből konstruáljuk a nyelvelsajátítás során, melyeket a beszéd fonetikai és nem fonetikai jegyeiből ismerünk fel annak rendszerszerű változatossága révén. Ebből pedig az következik, hogy az állandóság valójában a reprezentáció tulajdonsága, és (paradoxnak tűnő módon) éppen a variabilitásban rejlik. Az elmélet fontos következtetése tehát, hogy a variabilitás nem a beszéd folyamat másodlagos vagy mellékes, hanem lényegi, esszenciális jellemzője. Az elmélet fontos alapfogalma a **célkonfiguráció**, melyen az egyes beszédhangok idealizált artikulációs és akusztikai mintáját értjük, valamint a **célalulmúlás**, mely fogalom pedig a célkonfiguráció el nem érésének eseteit jelöli (ezt látjuk például a magyar felső és középső nyelvállású rövid magánhangzók esetében, vö. MÁDY 2008). Végezetül pedig kiemelt szerepet kap a modellben a **kontraszt**, ugyanis Lindblom javaslata szerint a kontraszt és a kontraszt erősítése szolgál a célalulmúlás révén „alulspecifikáltabb” beszédhangok közti különbségek megtartásának eszközéül (LINDBLOM 1990, LINDBLOM et al. 1992). A magánhangzók esetében a kontraszt erősítésének két prototipikus példája a túlartikulált beszédben megjelenő nagyobb magánhangzótér (vö. pl. WASSINK et al. 2007) vagy a fonológiai hosszúsági oppozíció erősítése szélsőségebb fizikai időtartamok (általában az alulartikulált esetben tapasztalhatóanál nagyobb időtartamú hosszú magánhangzók) realizációjának segítségével (vö. pl. TAJIMA et al. 2013).

1.1.2.2. A MAGÁNHANGZÓ-AZONOSÍTÁS KÍSÉRLETES EREDMÉNYEI ÉS A MAGÁNHANGZÓSZERŰ HANGOK FELDOLGOZÁSA A PERCEPCIÓS RENDSZERBEN

A magánhangzók (és magánhangzószerű hangok) azonosításáról máig viszonylag keveset tudunk, azonosításukról, az ahhoz szükséges (akusztikai) kulcsokról számos elmélet és kísérleti eredmény létezik. Mivel a magánhangzók legkarakterisztikusabb jegyei a hangzóspektrum prominenciái, azaz formánsai, így közvetve vagy közvetlenül a legtöbb hipotézis és kísérlet ezekre épít. Már az első, a magánhangzók akusztikai szerkezetét szisztematikusan vizsgáló kísérlet is összefüggéseket talált a magánhangzóspektrum prominenciái és a magánhangzók észlelése között (PETERSON–BARNEY 1952). Ezen is alapulva általánosan véve úgy véljük (és a legtöbb tankönyvben, kézikönyvben is azt találjuk), hogy a magánhangzók azonosítása szempontjából legfontosabb kulcs a magánhangzó első két formánsa (lásd pl. Gósy 2004). Azt azonban, hogy az első két formánssal kapcsolatosan pontosan milyen információt hasznosít a beszédpercepció, már nem fogalmazhatjuk meg ilyen egyszerűen.

A jelen alfejezetben két kísérletileg alátámasztott elméletet ismertetünk, valamint egy további, az ismertetett elméletekhez szorosan nem tartozó kutatási irányt, melynek az eredményeit a jelen értekezésben is hasznosítjuk. Az egyik elmélet a környéki hallórendszer működésével, azon belül is a hallóidegekkel és azok tüzelési mintázataival hozza összefüggésbe a magánhangzók azonosítását. A másik elmélet a spektrális prominenciák egymáshoz viszonyított, érzeti összefüggéseire alapoz. Végül azokat a vizsgálati eredményeket is ismertetjük, melyek arra világítanak rá, hogy a magánhangzók felismerésében nem csak a magánhangzó középső, viszonylag állandó (stacioner) szakasza játszik (vagy játszhat) szerepet, de a magánhangzók koartikulációjából származó dinamikus akusztikai információk is, melyek az adott magánhangzó és a szomszédos hangok egymás utáni ejtéséből származó, a fonémák határain átívelő, egymással átfedésben lévő artikulációs gesztusok lenyomatai.

BÉKÉSY (1960) kutatásai óta tudható, hogy egy elsődleges akusztikai elemzés, dekompozíció már a környéki hallórendszerben, azaz a fülben, egészen pontosan az abban található csigában is lezajlik, mégpedig az alaphártya működése révén. Az alaphártya frekvenciaérzékeny membrán, melyen a különböző (alacsonyabb és magasabb) frekvenciákra más-más szakaszok reagálnak, mégpedig fizikai kimozdulással. Ezt a kimozdulást az alaphártyához kapcsolódó szőrsejtek felfogják és idegi impulzussá alakítják, az impulzust pedig a hallóidegek továbbítják a központi idegrendszerbe, azaz az agyba. A hallóidegek azonban szintén frekvenciaérzékenyek, és saját karakterisztikus frekvenciával rendelkeznek. Ez azt jelenti, hogy minden idegrost egy bizonyos hangnyomás- és frekvenciatartományon tüzel, és minden idegrostra jellemző egy olyan tartomány, ahol az adott rost a legérzékenyebb (azaz ahol a válaszhoz szükséges küszöbérték a legalacsonyabb). Ez adja az idegrost hangolási görbét, illetve karakterisztikus frekvenciáját (STEVENS 1998; VICSÍ 2010a).

Az idegrostok tüzelési sajátosságain alapulva SRULOVICZ és GOLDSTEIN (1983) egy olyan elméletet dolgozott ki, mely szerint a komplex hangok (amilyenek például a magánhangzók

is) feldolgozásakor az idegek tüzelési mintázata az idegsejtek karakterisztikus frekvenciáinak függvényében szinkronizálódik a hangok spektrumának maximumaival (a magánhangzók esetében a formánsokkal), a komplex hangok feldolgozása tehát a spektrális maximumok és az idegsejtek sajátfrekvenciáinak összefüggéseitől függ. Későbbi pszichofizikai kísérletekben (lásd pl. DELGUTTE–KIANG 1984) ezt az elméletet a magánhangzók azonosításával és kategorizálásával kapcsolatosan közvetlenül is vizsgálták. A vizsgálatok tanúsága szerint az idegi kisülések mintázatában látható, hogy azok szinkronizálódnak a magánhangzóspektrumok F_1 - és F_2 -prominenciáival – míg más, nem beszéd típusú jel (például valamilyen széles sávú zaj esetén) az idegrostok a saját karakterisztikus frekvenciájukon tüzelnek. DELGUTTE és KIANG (1984) tehát egyrészt kimutatta, hogy a környéki idegrendszer azonosítja a magánhangzóspektrum prominenciáit, azaz a formánsokat, másrészt pedig kimutatta azt is, hogy (az előbbieknél is megfelelően) az idegrostok tüzelési mintázatai eltérőek azokra a magánhangzókra, amelyek a zárt-nyitott, illetve az elől-hátul képzettség dimenziók más-más pontján helyezkednek el. Ez azért lehetséges, mert – mint tudjuk – az elől képzett hangzók F_2 -értékei magasabbak a hátul képzettekénél, valamint a felső nyelvállású, zárt ejtésű hangok F_1 -értékei alacsonyabbak az alsó nyelvállású, nyílt hangzókénál, tehát az f_0 , az F_1 és az F_2 viszonya hangzóspezifikusan változik. A szerzők a vizsgálatukban azt találták, hogy ha a magánhangzóspektrum prominenciái megfelelően közel vannak egymáshoz (az f_0 - F_1 esetében ez a zárt képzésű magánhangzókra, míg az F_1 - F_2 esetében ez a hátul képzett magánhangzókra igaz), a prominenciák frekvenciái közötti szőrsejtek nem a saját karakterisztikus frekvenciájukon, hanem valamely hozzájuk frekvenciában közel eső prominencia frekvenciáján tüzelnek. Akkor viszont, ha a spektrum prominenciái a kritikus távolságnál nagyobb távolságra vannak egymástól, a két prominencia közti, köztes frekvenciákra érzékeny szőrsejtek a saját frekvenciájukon tüzelnek, tehát az adott prominenciákat (az f_0 -t az F_1 -től, illetve az F_1 -et az F_2 -től) elkülönítve érzékeljük. Mindez tehát azt jelenti, hogy bizonyos prominenciák (vagy prominenciapárok) közti viszonyok elkülönítik a zárt, felső nyelvállású hangzókat a nyílt, alsó nyelvállású hangzóktól, valamint a veláris hangzókat a palatálisoktól. DELGUTTE és KIANG (1984) magyarázata szerint az előbbit, azaz a zárttság érzetét az f_0 és az F_1 távolsága, míg az utóbbit, azaz az elől-hátul képzettség érzetét az F_1 és F_2 távolsága befolyásolja.

A magánhangzóspektrum és az idegrostok tüzelési sajátosságainak összefüggésein alapuló vizsgálatok tanúsága szerint tehát elmondható, hogy az elől és hátul képzett, valamint a zárt, felső nyelvállású és a nyitott, alsó nyelvállású magánhangzók már a környéki feldolgozási folyamatokban természetes kategóriákként különülnek el egymástól. Mindez azonban a nyíltság dimenziója mentén ilyen tisztán valójában csak a legalsó és legfelső nyelvállású hangzókra igaz, a köztes hangzók esetében ugyanis az idegek az alsó régiókban vagy az f_0 , vagy az F_1 frekvenciáján tüzelnek, az egyik prominencia tehát mindenképpen dominálja a tüzelési mintázatot. Ilyen értelemben tehát ezek a köztes hangzók csak a nyílt vagy legalsó nyelvállású magánhangzóktól különülnek el egyértelműen, melyek esetében az f_0 és az F_1 között vannak saját frekvencián tüzelő idegsejtek is (STEVENS 1998).

Az előbbiektől feltehetőleg nem független jelenségre és percepciós működésekre alapozza elméletét CHISTOVICH és LUBLINSKAJA (1979). Tanulmányukban pszichoakusztikai kísérletek során azt vizsgálták, hogy hogyan feleltethető meg egy két prominenciát tartalmazó akusztikus jel egy egy prominenciájú akusztikus jelnek. Ehhez a szerzők olyan (szintetizált) mintahangokat generáltak, amelyeknek két prominenciája, azaz formánsfrekvenciája volt, és az egyes stimulusokban ez a két prominencia szisztematikusan más-más távolságra helyezkedett el egymástól. A kísérletben részt vevők feladata az volt, hogy „keverjék ki” ezt a kétformános jelet egy egyformános jelből, melynek a spektrális maximumát (F') állíthatták. Az eredmények azt mutatták, hogy ha a mintahangban a két formáns egy kritikus távolságon belül helyezkedett el (egymáshoz képest), a jelet a tesztelők egyöntetűen egy olyan egyformános jellel képezték le, ahol az F' frekvenciaértéke a két formáns közötti, az amplitúdóval súlyozott átlagos frekvencia, azaz a „gravitációs középpont” (center of gravity) vagy súlypont frekvenciaértékét vette fel. Ha azonban a két formáns a kritikus távolságnál messzebbre került egymástól, az F' értékét a hallgatók (a formánsok relatív amplitúdójától is függően) vagy a mintajel F_1 -értékére, vagy a mintajel F_2 -értékére hangolták, tehát a két prominenciát külön-külön érzékelték. Így CHISTOVICH és LUBLINSKAJA arra következtetett, hogy egy kritikus távolságon belül, melynek értékét 3-3,5 Barkban⁵ állapították meg, a spektrális prominenciák az észleletben „összegződnek” (bár matematikai értelemben inkább kiátlagolódnak) egyetlen spektrális prominenciába, a két prominencia „gravitációs középpontjába”, súlypontjába vagy súlyozott átlagába.

Ugyanezt a jelenséget vizsgálta CARLSON munkatársaival (1970) ötfórmános szintetizált svéd magánhangzók segítségével, melyekhez a tesztelők egy kétformános jel második prominenciáját (tehát F_2 -jét) hangolhatták. A vizsgálat kimutatta, hogy az F_2 ' értékét a hallgatók általában a mintajel F_2 -értékére hangolták a veláris /a o u/ magánhangzók esetében (ahol az F_2 értéke relatíve alacsonyabb, és távol van az F_3 -tól). Ugyanakkor a palatális /e y ø ʉ æ/ magánhangzók esetében a hangolt F_2 ' értéke a mintajel F_2 és F_3 formánisa közti „gravitációs középpont” vagy súlypont frekvenciaértékét vette fel, a zárt, ajakréses (/i e/) hangzók esetében pedig az F_3 frekvenciaértékéhez közeli frekvenciát (az /e/ esetében az F_3 -nál valamivel alacsonyabb, míg az /i/ esetében az F_3 -nál valamivel magasabb értéket) kapott. A vizsgálat tehát azt mutatta ki, hogy ha az F_2 és F_3 formánsok távolsága kisebb volt, mint 3 Bark, ahogy a palatális magánhangzók esetében tapasztalható, a két formáns egyetlen, szélesebb sáv szélességű prominenciává kapcsolódott össze az észleletben, míg a veláris hangzók esetében, ahol az F_2 és F_3 távolsága nagyobb, mint a kritikus távolság, az F_2 -t különálló prominenciaként, pontosan azonosították a hallgatók. A szerzők percepciós tesztekkel azt is bizonyították, hogy a vizsgált svéd magánhangzók egyértelműen beazonosíthatók akkor, ha az első két formáns alapján szintetizálják őket (CARLSON et al. 1975), és arra is következtettek, hogy a magasabb formánsfrekvenciák feltehetően inkább az egyéni hangszínezet, és nem a magánhangzóhangszín meghatározása

⁵ A Bark a Hertz-értékkel szemben nem lineáris, hanem egy érzeti transzformáción alapuló akusztikai mértékegység. Erről és más nemlineáris mértékegységekről a fejezetben később még teszünk említést.

szempontjából fontosak. Ugyanakkor a vizsgálat fontos megfigyelése az is, hogy a palatális hangzók kétformásos közelítése futtatott percepciók tesztekben nem felelt meg olyan pontosan az ötformásos hangzók hangszínének, mint a velárisoké, ahol az F_2' az F_2 értékével egyező volt (STEVENS 1998 kiemelése). CARLSON és munkatársai (1970) eredményei tehát arra mutatnak, hogy a „gravitációs középponti” hatás szerepet játszik a palatális és veláris magánhangzók megkülönböztetésében, azaz annak az eldöntésében, hogy a megítélt hang hol helyezkedik el az elől és hátul képzettség dimenzióján. Ha ugyanis az F_2 és F_3 a kritikus 3-3,5 Bark távolságon belül van, a hang az észleletben palatális lesz, míg ha a két prominencia a kritikus távolságnál nagyobb távolságra van egymástól, a hangzót velárisként azonosítjuk.

Fontos azonban megjegyezni, hogy két spektrális prominencia egymáshoz képesti távolsága nem csak a „gravitációs középponti” hatás szempontjából fontos. Ha ugyanis két rezonancia frekvenciájában közelebb esik egymáshoz, az egyúttal mindkét rezonancia amplitúdónövekedését is okozza (vö. FANT 1960). Ezzel együtt viszont az is igaz, hogy ilyen esetekben egy, az e kettőtől távolabb eső harmadik rezonanciát relatíve alacsonyabb amplitúdó jellemez, azokhoz az esetekhez képest, amikor a rezonanciák távolsága kiegyenlített(ebb) a spektrális térben. Ha tehát a két legalacsonyabb frekvenciájú rezonancia, azaz az F_1 és az F_2 közel van egymáshoz, az egyúttal azt is jelenti, hogy az F_3 amplitúdója ezeknél jóval alacsonyabb – szemben azokkal az esetekkel, ahol az F_2 közelebb van az F_3 -hoz. Ez a megfigyelés STEVENS (1998) értelmezésében megerősíti azt az elképzelést, hogy a magánhangzó-minőség szempontjából inkább az első két formáns, és kevésbé az F_3 lehet döntő szerepű.

TRAUNMÜLLER (1981) a „gravitációs középponti” hatást az f_0 és az F_1 viszonylatában vizsgálta, mégpedig a magánhangzók zártságészleletének tekintetében. Eredményei szerint a kritikus távolság hatása az f_0 és az F_1 viszonyában is fontos, mégpedig azért, mert a felső nyelvállású, alacsony F_1 értékkel rendelkező hangok spektrumát az észleletben egyetlen (az F_1 értékének megfelelő) frekvenciakomponens dominálja, míg a nyíltabb ejtésű, magasabb F_1 értékkel rendelkező hangokét ugyanezen a tartományon két prominencia, melyek az f_0 és az F_1 frekvenciáinak felelnek meg. Következésképpen TRAUNMÜLLER konklúziója az, hogy az f_0 és F_1 között lévő tonotópiás távolság⁶ egyértelműen befolyásolja a zártságészleletet.⁷ Ezt az elméletet SYRDAL és GOPAL később egy produkciós vizsgálatában is igazolja, megmutatva azt, hogy a felső nyelvállású magánhangzók esetén az f_0 és az F_1 általánosan 3 Barknál kisebb távolságra esik egymástól, függetlenül attól, hogy ezek milyen frekvenciaértéken realizálódnak. Ez azt jelenti, hogy SYRDAL és GOPAL eredményei szerint az f_0 és F_1 távolsága kisebb 3-3,5 Barknál mind a gyermekek, mind a felnőtt férfiak, mind pedig a felnőtt nők ejtésében (SYRDAL–GOPAL 1986).

⁶ A tonotópiás szerveződés azt jelenti, hogy a különböző frekvenciákra érzékeny területek különböző területeken lokalizálódnak, ráadásul az alaphártya esetében ez a szerveződés igen szisztematikus is: a szomszédos frekvenciákra érzékeny területek egymás mellett találhatóak (vö. KÁLLAI et al. 2008).

⁷ TRAUNMÜLLER szándékosan a „zártság” terminust használja a felső vagy alsó nyelvállású hangzók helyett, mert terminológiájával nem a nyelvállásfokra, hanem az állkapocsnyitás nagyságára kíván utalni, ami vélekedése és az általa hivatkozott szakirodalmi tételek szerint jól korrelál a percepciók „zártsággal”.

A zártság észleletéhez ugyanakkor valószínűleg további akusztikai jellemzők is kapcsolhatók. Már a magánhangzók akusztikai szerkezetét először szisztematikusan megvizsgáló PETERSON és BARNEY (1952) is megfigyelte például azt, hogy (felolvasásban, izolált szavakban) a magasabb nyelvvállásfokon képzett, zártabb hangzók f_0 -értéke is magasabb, mint az alacsonyabb nyelvvállásfokon, nyíltabban képzett hangzóké (az f_0 értéke tehát valamelyest magánhangzófüggőnek tekinthető). Ez azt jelenti, hogy a magasabb nyelvvállású hangok f_0 - és F_1 -értékei közti különbség nem csak az alacsony F_1 miatt kicsi, hanem a magasabb f_0 -érték miatt is, tehát feltételezhetjük, hogy a zártabb hangzók esetében a magánhangzófüggő f_0 értékek is erősítik a zártságészleletet.

Az itt említett elméletekben, kísérletekben fontos közös vonás, hogy ezek a magánhangzók észlelését, illetve a magánhangzó-kategóriák elkülönülését nem a lineáris vagy abszolút fizikai értékeken alapuló adatok mentén, hanem a lineáris adatok érzeti transzformáltjaival összefüggésben vizsgálják, ilyen értelemben tehát a magánhangzók variábilis megvalósulásainak kérdését valóban a percepció felől közelítik meg. De hogy ez miért hangsúlyozandó? A beszédpercepció vizsgálatának egyik első kérdése szerint – melyet PETERSON és BARNEY (1952) eredményei nyomán fogalmazott meg a szakirodalom – a percepciós folyamat leírásakor azt is meg kell tudnunk válaszolni, hogy hogyan dolgozódnak fel a percepciós rendszerben a különböző magánhangzók megvalósulásai között tapasztalható nagymértékű átfedések, azaz az a tény, hogy az egyes kategóriák határai nem élesek a produkcióban. Ha ugyanis a magánhangzókat a lineáris Hertz-értékek szerinti $F_1 \times F_2$ magánhangzótérben kezeljük és ábrázoljuk (ahogyan azt az idézett szerzők is tették), akkor az egyes magánhangzó-kategóriák megvalósulásait felölélő sűrűsödési ellipszisek, „szigetek” valóban jelentős átfedésben vannak. Ha azonban az akusztikus magánhangzótér ábrázolásához valamely nemlineáris, frekvenciavetemítésen alapuló, érzetileg transzformált skálák mentén felvett kétdimenziós projekciót alkalmazunk (például logaritmikus skálára hozzuk, illetve mel- vagy Bark-értékekbe transzformáljuk a Hertz-értékeket),⁸ a lineáris adatok alapján felvett magánhangzótérben tapasztalható átfedések csökkennek (vö. SYRDAL–GOPAL 1986; NEARY 1989) – így a PETERSON és BARNEY (1952) eredményei nyomán megfogalmazott probléma jelentősége is csökken.

A beszédpercepció szakirodalmának egyik alapvetően elfogadott nézete szerint a koartikulációs hangátmenetekben kódolt, a magánhangzók képzésére is vonatkozó akusztikai információknak köszönhetően a magánhangzók nagyobb arányban ismerhetők fel, ha a hallgatók mássalhangzó-környezetben ejtve hallgatják őket. STRANGE és munkatársai (1976) /p/ /p/ hangkörnyezetbe ágyazott magánhangzók vizsgálatokor azt találták, hogy míg a mássalhangzó-környezetben ejtett magánhangzók azonosítása 80-90% körül mozgott, az izolált

⁸ Míg a logaritmikus transzformáció során a frekvenciaértékeket zenei félhangokká transzformáljuk, a Bark-skála az auditív maszkoláson alapszik, és az ahhoz kapcsolódó hallási eseményeket modellező szűrőket használ a transzformációhoz. Ez utóbbit, tehát a Bark-szűrősort használó transzformációt pszichoakusztikai kísérletekkel dolgozták ki. A logaritmikus és a Bark-szűrősorton alapuló transzformáció közti különbség valójában nem nagy, és főképp az 500 Hz alatti tartományt érinti. A különbség abban áll, hogy 500 Hz alatt a Bark-skála a logaritmikus skálánál lineárisabb viszonyban függ össze a Hertz-értékekkel, ez pedig realisabban képezi le a hallási érzékenységet (NEARY 1989).

ejtésű hangzókat 30-40%-ban tévesztették más beszédhangokra a hallgatók, annak ellenére, hogy formánsértékeiket tekintve az izolált ejtésű hangzók a mássalhangzó-környezetbe ágyazott megvalósulásokhoz képest túlarguláltak voltak, tehát egymástól inkább eltérőek, kontrasztosabbak. Ugyanezen tanulmány második kísérletéből az is kiderült, hogy a mássalhangzók pozitív hatása még akkor is kimutatható, ha a mássalhangzók minősége nem megjósolható (tehát bemutatásonként változik). JENKINS és munkatársai (1983) később azt is kimutatták, hogy a formánsátmenetekben rejlő akusztikai információ még a magánhangzó mediális (tiszta fázisának tekinthető) 50-60%-ának törlése mellett is nagyon magas arányban megőrzi a magánhangzó-minőségre vonatkozó információt, ugyanis nem tapasztaltak különbséget a közepén törölt szótagra és a teljes szótagra mutatkozó azonosítási arányok között. A magánhangzó és az őt követő vagy megelőző mássalhangzó közti formánsátmenetek, azaz a koartikuláció akusztikai vetülete tehát úgy tűnik, hogy nagyon fontos dinamikus akusztikai kulcsokat jelent a magánhangzók azonosításában.

NEARY (1989) az akusztikai információ négy típusát különíti el mint az észlelésben (legalábbis az angol magánhangzók észlelésében) fontos, de változó súlyú kulcsokat. Az első a) típusba a statikus tulajdonságok tartoznak, mint például a magánhangzó középső, nagyjából az időben állandónak tekinthető szerkezetű szeletének (ezt magyarul tiszta fázisnak is nevezik, lásd pl. GÓSY 2004) formánsfrekvenciái és alapfrekvenciája. A második b) típusba Neary a dinamikus tulajdonságokat sorolja, ilyenek a magánhangzóban rejlő inherens spektrális változások és a mássalhangzó-környezet hatására megjelenő formánsátmenetek, melyekkel kapcsolatos kísérletes eredményekről az imént tettünk említést. A harmadik c) csoport az intrinzikus, azaz a szegmentumon belüli frekvenciarelációkból fakadó tulajdonságok csoportja, ide különösen a formánsok egymáshoz és az alapfrekvenciához viszonyított relációja tartozik. A negyedik d) csoportba az extrinzikus, azaz a transzsegmentális relációs tulajdonságok tartoznak, mint az adott magánhangzónak az adott beszélő többi magánhangzójához viszonyított relatív időtartama, valamint a magánhangzó a többi magánhangzó formánsértékeihez viszonyított relatív frekvenciaszerkezete – ezek az információk ugyanis viszonyítási pontokként segítik a percepció „kalibrálását” a beszélő egyéni sajátosságaiból adódó hatások eliminálása (azaz a beszélőnormalizálás) révén (erről lásd még pl. LADEFOGED–BROADBENT 1957).

Mindabból, amit a beszédhangok és különösen a magánhangzók észleléséről eddig tudunk, és amely tudásnak egy szeletét a jelen alfejezetben is igyekeztünk – a teljesség igénye nélkül – körüljárni, fontosnak tartjuk kiemelten hangsúlyozni a következőket. A számos kísérletes munka eredményeképpen ma már sejthető, hogy melyek azok az információk, amelyek elegendőek a magánhangzó-azonosításhoz, illetve melyek azok, amelyeket a percepció működése a magánhangzó azonosításában hasznosítani tudnak. A normál módú beszédben minden bizonnyal kiemelt szerepe van az F_1 és F_2 formánsoknak, valamint (főként az F_2 -vel való auditív interakciója miatt) az F_3 formánsnak a palatálisok esetében, továbbá az f_0 -nak, azaz az alapfrekvenciának. Úgy tűnik ugyanis, hogy ezek az akusztikai paraméterek és ezek viszonyai elegendő információt hordoznak a magánhangzók minőségének beazonosításához.

Azt azonban, hogy ezekből az információkból mi elengedhetetlen a beszédpercepció számára, hogy mi történik akkor, ha az alapfrekvencia emelésének következtében ezek az információk nem elérhetők (vö. az előző fejezetben említett alul-mintavételezettség fennállása esetén), vagy hogy ezeket az információkat pontosan milyen módon, milyen súlyozással hasznosítja az emberi percepció, nem tudjuk biztosan.

1.1.2.3. A MAGYAR MAGÁNHANGZÓK AZONOSÍTÁSÁNAK KÍSÉRLETES EREDMÉNYEI

A magyar magánhangzók azonosítását, egészen pontosan a frekvenciaszerkezet hatását a magyar magánhangzók minőségének felismerésére GÓSY Mária (1989) vizsgálta behatóbban egy három fázisból álló kísérletsorozatban.

Az első fázisban GÓSY (1989) a 14 magyar magánhangzót egy férfi beszélő bemondásában rögzítve, izolált ejtésben azonosította a hallgatókkal, mégpedig úgy, hogy a magánhangzókat kilencféle különböző felső határértékre beállított alul áteresztő szűrővel szűrte. A kísérlet több tesztből állt, melyek során a szűrő felső határának frekvenciaértékét fokozatosan növelték, így az áteresztett tartomány, tehát az áteresztett akusztikai információ mennyisége is fokozatosan nőtt a kísérlet előrehaladtával. A vizsgálatban GÓSY – többek közt – az alábbi eredményekre jutott.

Egyes hangzók (elsősorban az /o/) esetében már az F_1 megjelenése előtt (tehát akkor, amikor az F_1 frekvenciaértéke magasabb volt, mint a szűrő felső határának értéke) lehetséges volt a magánhangzók helyes azonosítása. Az /o/ esetében – bár szükséges volt az F_1 , – az F_2 megjelenése előtt (azaz az F_2 alatti frekvenciasáv alapján) is biztosítottnak tűnt a magasabb arányú helyes azonosítás. Az eredmények ugyanakkor arra is mutattak példát, hogy bizonyos hangok esetében nem hogy javított, hanem inkább még rontott is az azonosítási arányokon az egyre gazdagabb akusztikai információ: míg az /i/ a 390 Hz felső határértékkel rendelkező frekvenciasávig (azaz a csak az F_1 -et tartalmazó spektrumú hangrészlettel) 80%-ban azonosíthatónak bizonyult, az egyre növelt frekvenciájú felső határ egyre alacsonyabb azonosítási arányokhoz vezetett. Ezen adatok alapján a szerző azt a megállapítást tette, hogy az F_1 -et tartalmazó frekvenciasáv valamilyen formában tartalmaz az F_2 -re vonatkozó információt is, valamint azt, hogy az F_1 és F_2 közé eső tartomány nem erősített frekvenciakomponensei is fontos szerepet játszanak a magánhangzó-minőségek felismerésében. Bizonyos értelemben tehát GÓSY (1989) hasonló következtetésre jutott, mint néhány korábban idézett vizsgálat lefolytatói, hiszen SRULOVICZ és GOLDSTEIN (1983), valamint DELGUTTE és KIANG (1984) vizsgálataikban szintén hangsúlyozzák az F_1 és F_2 közti frekvenciatartomány fontosságát. Emlékeztetőül, az idézett kutatók megállapítása szerint a hallóidegek tüzelési mintázatai alapján az egymáshoz adott (3-3,5 Barknál kisebb) távolságra elhelyezkedő F_1 és F_2 rezonanciák „összeolvadnak”, míg az egymástól ennél nagyobb távolságra elhelyezkedő rezonanciák elkülönülnek az észleletben, ilyen értelemben tehát az F_1 és F_2 közti frekvenciatartomány megléte és kiterjedése elválasztja egymástól a magánhangzók egyes csoportjait.

A tévesztésként megjelenő beszédhangok zömét a legalacsonyabb felső határértékkel szűrt (tehát legkeskenyebb sávon áteresztett) hangmintákon GÓSY (1989) összefoglalása szerint elsősorban zárt vagy felső nyelvvállású, hátul képzett hangok képezték. Ez a megfigyelés összefüggésbe hozható TRAUMÜLLER (1981) fentebb idézett eredményeivel. Emlékeztetőül, TRAUMÜLLER (1981) vizsgálatában azt találta, hogy a felső nyelvvállás hangzók észleleti lenyomatában az f_0 és F_1 tartományán mindössze egyetlen észleleti prominencia van jelen valahol az f_0 , illetve az F_1 frekvenciahelyének környékén, míg a nyíltabb hangzók esetében kettő, mégpedig az f_0 és az F_1 frekvenciáján. Ez alapján feltehető, hogy ha az adatközlőknek olyan akusztikai jelet kell beszédhangként azonosítaniuk, amelyben formánsinformációk híján mindösszesen az alaphang jelenik meg prominenciaként, azt inkább zártabb hangzóként azonosítják – és valóban, GÓSY (1989) eredményei éppen erre a feltételezésre szolgálnak bizonyítékkal.

GÓSY (1989) a kísérlet második fázisában sáváteresztő szűrőket alkalmazott az izoláltan ejtett magánhangzók akusztikai módosítására, melynek felső határa minden esetben 2700 Hz volt, alsó határát azonban a szerző fokozatosan állította 270 Hz-től indulva 2200 Hz-ig (így a szűrés után létrejött hangzók sáv szélessége a fokozatokkal egyre csökkent). A kísérlet e fázisának legfőbb tanulságaként GÓSY (1989) azt a következtetést vonta le, hogy a sávszűrt magánhangzók azonosítása a vizsgált sávokon belül végig pontos maradt, tehát az egyes hangzók azonosítása még akkor is viszonylag magas arányú volt, amikor az F_1 és F_2 formáns gyakorlatilag teljesen eltűnt (ki volt szűrve) az akusztikai jelből, és a szűrő által áteresztett tartományba csak az F_3 esett. (A szerző megjegyzi ugyanakkor azt is, hogy ez alól a tendencia alól egyes hangzók bizonyos kondíciókban kivételt képeztek.) Az eredmények további taglalásakor a szerző arra is kitér, hogy a legmagasabb alsó határértékkel szűrt tartományokban inkább az elől képzett hangzók azonosítása látszott pontosnak, még akkor is, ha azok F_2 -je a kiszűrt tartományba esett, azaz ha az áteresztett tartományban csak az F_3 és nem erősített frekvenciahelyek maradtak. Ebből a szerző az F_3 kiemelt jelentőségére következtetett: vélekedése szerint az F_3 jellemző és meghatározó a magánhangzó-minőség szempontjából, mégis inkább csak másodlagos felismerési kulcsnak tekinthető, hiszen az F_1 és F_2 jelenlétében pusztán az F_3 manipulálása nem befolyásolta a fonémadöntést, azok hiányában azonban az F_3 jelentősége megnőtt. Ismét csak korábbi vizsgálatokkal párhuzamot vonva megemlítendő, hogy GÓSY (1989) konklúziója összhangban van CARLSON és munkatársainak (1970) eredményeivel. Azok szerint ugyanis a palatálisoknál az F_2 és F_3 percepciók közelsége miatt az F_2 és F_3 helyén valójában nem két, hanem csak egy, frekvenciájában a két formáns közé eső érzeti prominencia jelenik meg – ilyen értelemben tehát az F_3 az F_2 -vel „összefonódva” képez egyetlen akusztikai kulcsot a palatálisok azonosításában.

A vizsgálat harmadik fázisában GÓSY (1989) felül áteresztő szűrő segítségével manipulálta a beszédhangokat, mégpedig úgy, hogy a szűrő alsó határát módosította 1000 Hz és 3300 Hz között. Az eredmények szerint a legmagasabb alsó határértékkel (tehát a legkeskenyebb sávot átengedő szűrővel) szűrt hangminták közül a legmagasabb arányban az /i i:/ és /y y:/ magánhangzókat azonosították a hallgatók, melyek második formánsa egyébként a legmagasabb

volt a kísérletben. Az /i i:/ és /y y:/ közti különbséget, azaz az ajakkerekítésre vonatkozó információt a szerző, konklúziói szerint, két akusztikai információban látja kulcsoltnak. Az egyik tényező a részösszetevők eltérő intenzitása, a másik pedig az az ismert összefüggés, hogy az ajakréses hangzók intenzitása általában véve nagyobb, mint az ajakkerekítéseseké (főként azért, mert az ajakkerekítésesek esetében nagyobb az ajaknyílás-felület, mely egységnyi idő alatt nagyobb levegőmennyiséget, de következésképpen kisebb teljesítménnyel ereszt át).

Egy következő vizsgálatban Gósy (1989) az időtartam és a frekvencia-összetevők szerepét vizsgálta szintetizált magánhangzók azonosításában. Ebben a percepciókísérletben a résztvevőknek különböző frekvenciaszerkezetű és különböző időtartamú magánhangzókat kellett azonosítani. Az eredmények azt mutatták, hogy ahogyan a magánhangzók időtartama növekedett, úgy a hallgatók az F_1 -értékek egyre szűkebb tartományában azonosítottak egy adott hangzóminőséget, az időtartam növekedésével ugyanis inkább nyíltabb magánhangzókat véltek felismerni (pl. /u/ ~ /u:/ helyett /o/ ~ /o:/-t). (Megjegyzendő, hogy ez alól a tendencia alól kivételt képeztek a labiális palatális /y/ ~ /y:/ /ø/ ~ /ø:/ hangzók, valamint az is, hogy a bemutatott adatok szerint nem fordult elő, hogy egy magánhangzó az észleletben ne egy szomszédos kategóriára „ugrott” volna.) A kísérlet tanulságát a szerző úgy foglalja össze, hogy a magánhangzó időtartama fontos kulcs a magánhangzó-minőség felismerésében: bár azokban a helyzetekben, amikor a szintetizált magánhangzók formánsértékei a magyar magánhangzókra legjellemzőbb formánsfrekvenciák értékeit vették fel, az időtartam mint akusztikai kulcs másodlagosnak bizonyult, más esetekben az időtartam a frekvenciaszerkezettel szimultán, elsődleges kulcsnak mutatkozott.

Gósy (1989) kísérleteit később főként olyan magánhangzó-azonosítási vizsgálatok követték, melyek a magánhangzók fizikai és nyelvi időtartamának összefüggéseit, valamint a nyelvi hosszúság és a frekvenciaszerkezet összefüggéseit elemezték elsősorban az elméletileg csak fizikai időtartamukban különböző felső és középső nyelvállású párok esetében (pl. Kovács 1998; Mády–Reichel 2007). Ezen vizsgálatok bemutatására a jelen értekezés keretei között azonban nincs módunk, hiszen – bár a magyar magánhangzórendszert érintő fontos kérdéseket vizsgálnak auditív szempontok szerint –, a jelen értekezésben ismertető kísérletek felől nézve a fonológiai oppozíció tagjainak megvalósulására és auditív oppozíciójára vonatkozó eredmények nem relevánsak. Bemutatandó kísérleteink ugyanis (az énekelt magánhangzók azonosítására fókuszálva) e párok tagjai közül minden esetben csak a hosszú magánhangzókat vizsgálják megközelítőleg egységesen hosszú időtartamban (kitartott ejtésben), mert elsősorban nem a magánhangzók észlelt nyelvi hosszúságának, hanem minőségének változásait kívántuk elemezni az éneklésben (vö. a 2.-től 7. fejezetek *Anyag, módszer és kísérleti személyek* alfejezeteiben található leírásokkal).

Összefoglalóan elmondhatjuk tehát, hogy a magyar magánhangzók azonosításában a nemzetközi szakirodalomban tapasztalt bizonyos tendenciákat már kísérletes úton is igazolták: megerősítést nyert az F_1 és F_2 formánsok kiemelt szerepe a hangzóminőség perceptuális kialakításában, illetve azt is láttuk, hogy nem feltétlenül szükséges mindkét formáns teljes sávzélességű megléte a magasabb arányú azonosításhoz.

A jelen kötet vizsgálatának középpontjában egy sajátos beszédképzési móddal létrehozott magánhangzók állnak: az énekelve ejtett, vagyis az énekelt magánhangzók. A témamegjelölés azt sugallja, hogy ez a kérdéskör elkülönül a normál beszédmódban ejtett magánhangzók produkciójának és percepciójának problémájától – a következő alfejezet célja az, hogy az éneklésről eddig ismeretes artikulációs, akusztikai és percepciósi sajátosságok tárgyalásán keresztül megvilágítsuk, ez valóban így van. A következőkben tehát az énekelt magánhangzók fonetikai szempontú szakirodalmának a lehetőségek szerint alapos, de minden bizonynyal nem teljes összefoglalására törekszünk, figyelmünket elsősorban az énekelt magánhangzók felismerésének kérdése szempontjából fontosabb aspektusokra fókuszálva.

1.2. Az éneklés és az énekelt magánhangzók fonetikai jellemzése

Az éneklés fonetikai szempontokat érvényesítő szakirodalmának összefoglalása sem kezdődhet másként, mint az „éneklés” fogalmának pontos meghatározásával. Mivel azonban e kötet nem kíván szólni sem zenetörténeti, sem pedig zeneelméleti kérdésekről, a fogalom meghatározásában sem törekszünk (és nem is törekedhetünk) a zenetudomány felől nézve maradéktalanul pontos meghatározásra. E „definíció” célja sokkal inkább az egyértelmű és pontos beazonosíthatóság kell, hogy legyen. Ennek fényében úgy fogalmazhatunk, éneklésen a jelen értekezésben a tradicionális vagy nyugati típusú operai műfajt értjük, mi több, a jelen értekezésben e műfajon belül is elsőként a magas alaphangfrekvenciájú szoprán éneklésre fókuszálunk.

Az opera hagyományosan igen sajátos műfaj – bizonyos értelemben mind az igen korai, mind pedig a kortárs ének műfajokkal szemben áll. A mai zenei műfajokhoz hasonlóan ugyanis nagy hangerőt igényel – hiszen az operai szólóénekeseknek egész zenekarokat kell „túlénekelnie” –, mégis énekesei saját beszédképző szerveiken kívül hagyományosan nem használnak semmilyen (külső) hangerősítést.

Az olasz opera megjelenésekor a reneszánsz világ polifonikus zenéjében különleges újdonságot jelentett, hiszen alapvetően szólisztikus műfaj. Európai elterjedésekor, a 17. század környékén a műfajt még elsősorban az igen lágy hangjukról, finom díszítésekkel jellemezhető stílusukról ismert kasztrált énekesek (castrati) dominálták. Az ekkoriban lezajló zenei fejlődés, a zenei műfajok változása azonban a használt hangszerek felépítésének változásaival is együtt járt, az egyre jobb rezonátorképességekkel megépített hangszerek hangereje egyre növekedhetett, így a halkabb szólisták ideje hamar leáldozott. A hangszerek felépítésének megváltozása mellett ráadásul a kísérőzenekar mérete is gyarapodott, együtt a zenei eseményeknek helyt adó terek növekedésével. Mindennek eredményeképpen tehát a 19. századra nem csak az énekhang minőségének, de „mennyiségének” is módosulnia kellett, hogy az egyre növekvő mennyiségű auditív maszkoló (azaz elfedő, a hallhatóságot akadályozó) jelenségeket

leküzdje. Így alakultak ki egyes új hangminőségek (a koloratúr- és a drámai szoprán), alakultak át az egyes szerepek (egyre nagyobb szerepet kaptak a magas alapfrekvenciákon éneklő énekesek, a tenorok, baritonok vagy a mezzoszopránok), elkezdett egyre feljebb és feljebb terjeszkedni a tenor és szoprán hangterjedelem. Így jött létre tehát (nagyjából az 1800-as évek végére) a ma is ismert operai hangzás és énektechnika. Az énekhang minősége vagy az énekhang hangszínezete azonban nem pontosan definiált, a technika örökítésében alapvetően a mai napig elsősorban egy-egy tanár reprezentálja a stílusbeli és hangzásbeli kívánalmakat (MITCHELL 2005).

A ma ismert operaéneklés hagyománya tehát többszáz éves. Az operai énekhang fontos kritériuma a nagy hangerő és a lehetőség szerint a teljes alapfrekvencia-tartományon homogén hangszínezet. Bár ma már rendelkezésre áll a hangerősség növelésére számos (elektro)technikai megoldás, a szigorúan vett operai műfaj (szemben például a musical vagy operett műfajokkal) nem él ezekkel a lehetőségekkel – a hagyományos kritériumoknak megfelelő hangszínezetű és hangerejű hangadás kulcsa pusztán az énekes rezonátorüregeinek, egészen pontosan az énekes toldalékcsovi rezonanciáinak jobb, hatékonyabb kihasználásában rejlik.

A jelen értekezés témája kizárólag a magas alapfrekvenciájú szoprán éneklés. Ennek oka, hogy vizsgálatok tanúsága szerint a magasabb frekvenciák énekesei – de elsősorban a szopránok – az alacsonyabb hangosztályok (basszus, bariton vagy alt) énekeseihez képest egészen más jellegű akusztikai problémák megoldására kényszerülnek, melyeket (értelemszerűen) erősen eltérő technikával, eltérő stratégiákkal oldanak meg. Ezen túlmenően a jelen vizsgálatok elsősorban csak a magánhangzókra fókuszálnak, hiszen az énekes bizonyos értelemben csak e hangok segítségével „énekel”: mivel a magánhangzók képzésekor a levegő áramlása akadálytalan, és a magánhangzós részek nem tartalmaznak zörejes elemeket, csak a magánhangzók képzését általánosan jellemző zöngét, így ezek a hangzók a legalkalmasabbak az alaphangélmény előidézésére, tehát a dallam kifejezésére. Valószínűleg ennek köszönhető az is, hogy bizonyos értelemben alapvetően a magánhangzók, a magánhangzók dominanciáján alapszik a *bel canto* néven is ismeretes operai stílus, így a jelen kötetben mi is ezeket a beszédhangokat vesszük górcső alá.

A magas alapfrekvenciájú éneklés problematikus feladatnak tekinthető mind a hangerő, mind a hangszínezet, mind pedig a hangszín, azaz az egyes magánhangzó-minőségek szempontjából, hiszen ezek a fonetikai tulajdonságok változnak a beszéd- vagy énekhang alapfrekvenciájának változtatásával, legfőképpen az f_0 extrém emelésével. Érdekes módon azonban ezek a nehézségek – úgy tűnik – kivétel nélkül ugyanazon fonetikai problémára vezethetők vissza, mégpedig arra, hogy a szoprán énekesek gyakran kényszerülnek olyan magas alapfrekvenciájú hangok kiéneklésére, melyek a beszédhang minősége szempontjából kiemelten fontosnak tartott első formáns frekvenciájánál magasabbak. Ha az énekes „túllőné” az F_1 -et, azaz magasabb zenei hangon énekelne, mint az adott magánhangzónak a beszédben jellemző F_1 -értéke, és az adott magánhangzót emellett változatlanul, a beszédben megszokott módon és F_1 értékkel (azaz ugyanolyan nyíltsági fokon) próbálná kiejteni, az egyrészt a kiadott hang jelentős hangerőbeli csökkenését okozná, másrészt pedig igen erős hangszíne-

zetbeli váltást is, melyek közül ebben a műfajban egyik sem kívánatos (vö. pl. SUNDBERG 1987). Ráadásul az F_1 „tüllövése” egyúttal azt is jelentené, hogy az F_1 egyszerűen eltűnne a magánhangzóspektrumból, ami – az eddig taglalt ismereteink alapján – feltehetően az ejteni szándékozott magánhangzó felismerhetőségének komoly csökkenését okozná. Egy operaária meghallgatása során azonban a fentieket nem vagy nem maradéktalanul tapasztaljuk, képzett énekesek esetében ugyanis ezek a problémák nem vagy legalábbis csak részben állnak fenn. A professzionális énekesekre alapvetően nem jellemzők a hirtelen hangszínezetbeli váltások, miközben énekhangjukat nem nyomja el a zenekari kíséret hangja sem. Ugyanakkor a szopránok ejtette magánhangzók néha bizonytalan minőségét és az énekelt szövegek értési nehézségeit feltehetően mindannyian tapasztaltuk már. Kérdésként merül fel tehát, hogy mit tesz az énekes annak érdekében, hogy hangja homogén hangszínezetű és nagy hangerőségű maradjon magasabb alapfrekvenciákon is, és mi történik ennek következtében az ejteni szándékozott magánhangzók, a magánhangzók minőségével, valamint hogyan alakul ezen ejteni szándékozott magánhangzók felismerhetősége.

A magas alapfrekvencián énekelt magánhangzókat vizsgáló szakirodalomban ma már szinte közhelynek számít az a feltételezés, mely szerint azokban a helyzetekben, amikor az F_1 értéke alacsonyabb lenne, mint az éneklendő alaphang, azaz az f_0 frekvenciaértéke (ezen akusztikai helyzet rövid jelölése: $F_1 < f_0$), az énekes az F_1 értékét a megemelt f_0 értékére vagy valamivel a fölé hangolja (vö. pl. SUNDBERG 1987).⁹

SUNDBERG (1979 és 1987) szerint (aki a szoprán énekesek alapfrekvencia-függő produktív tendenciáit elsőként tanulmányozta mélyrehatóan, és akinek vizsgálatairól a következőkben még sok szó lesz) ez az $F_1 : f_0$ hangolásról szóló feltételezés akár empirikus bizonyítékok nélkül is igen stabil lábakon állhat. Érvélese szerint ugyanis már pusztán az F_1 -nek az f_0 -ra hangolásával, azaz a megszólaló hang alapfrekvenciájának felerősítésével 20 vagy akár 30 dB hangnyomásszint-növekedés (azaz hangerő-növekedés) is elérhető lehet. A formánshangolás, azaz a toldalékcső rezonanciatevékenységének e jobb kihasználása ráadásul úgy idézi elő a hangerő-növekedést, hogy ahhoz semmilyen további vokális erőfeszítés (pl. a szubglottális nyomás növelése, azaz kiabálás) nem szükséges (lásd még TITZE 1994, 2004 és 2008).

Az F_1 -nek az f_0 -ra hangolásáról szóló felvetés empirikus úton igen nehezen igazolható, hiszen a magas alapfrekvenciákon ejtett magánhangzók akusztikai lenyomatában (ahogyan azt az 1.1.1.2. *alfejezetben* bemutatuk) a spektrális alul-mintavételezettség jelensége miatt a hagyományos, Fourier-transzformáción alapuló elemzések segítségével nem határozhatók meg megbízhatóan az egyes formánsok középértékei, így az F_1 frekvenciája sem. Ennek eredményeként SUNDBERG első vizsgálatait – és az $F_1 : f_0$ hangolás akusztikai stratégiájának első, elméleti felvetését – számos további vizsgálat követte, melyek vagy közvetett, artikulációs

⁹ Itt jegyezzük meg, hogy a jelen értekezésben a formánshangolási stratégiákat a nemzetközi szakirodalom (lásd pl. JOLIVEAU et al. 2004; GARNIER et al. 2010) egyik elfogadott konvenciójának megfelelően kettősponttal jelöljük olyan módon, hogy a kettőspont bal oldalán a hangolt formánst, a jobb oldalán pedig a hangolási célt tüntetjük fel, például: $F_1 : f_0$, melynek jelentése 'az F_1 -nek az f_0 -ra hangolása'.

adatokból derivált akusztikai adatok alapján, vagy közvetlen, valamely újszerű vagy a tudományterületen addig nem elterjedt mérési módszertan alkalmazásával nyert akusztikai adatokból igyekeztek meghatározni az énekelt magánhangzók ejtésekor jellemző rezonanciafrekvenciákat, elsősorban az F_1 és az F_2 értékét. E vizsgálatok fő kérdése az volt, hogy milyen módokon képes a magas alapfrekvencián éneklő szoprán énekes rezonátorüregeinek jobb kihasználására, így nagyobb hangerő és szép, homogén hangszínezet létrehozására, továbbá – egyes kísérletekben – az a kérdés is felmerült, hogy miként realizálódnak ezek a módosítások az egyes magánhangzó-minőségek függvényében.

A formánshangolás azonban természetesen a hangerő-növekedés előidézése mellett az egyes magánhangzó-minőségek produkciójára, az egyes minőségek közti különbségek fenntartására is hatással van. Az $F_1 : f_0$ hangolás eredményeként ugyanis a magasabb alapfrekvenciákon az egyes magánhangzók akusztikai szerkezete (a kategóriákon átívelő módon egységessé váló f_0 és F_1 érték miatt) egymáshoz a beszédben tapasztalhatónál jóval hasonlóbb lesz, ami valószínűsíthető módon a perceptuális különbségek csökkenését is maga után vonja. Ebből a feltételezésből következően az énekelt magánhangzók vizsgálatának másik nagy területe a magánhangzók percepciójának kutatása. Eddig ezen a területen belül elsősorban az azonosíthatóság határait és segítő körülményeit kutatták.

Mivel az énekelt és a beszélve ejtett magánhangzók közti különbségek jóval nagyobbak is lehetnek, mint a beszédben egyébként is tapasztalható (beszélőn belüli és beszélők közötti) variabilitás,¹⁰ az énekelt magánhangzók észlelésének vizsgálata egyúttal a beszédpercepció működéséről alkotott képünket is tovább árnyalhatja, hiszen ennek révén az észlelés rugalmassága vagy akár korlátai is jobban megfigyelhetők. Az énekelt magánhangzók produkálhatóságának és azonosíthatóságának feltételeiről és korlátairól szóló ismeretek a gyakorlatban is hasznosíthatók lehetnek: felhasználhatók az énektechnika tanításában, az operaszövegek fordításában, vagy figyelembe vehetők akkor, amikor egy-egy opera-előadás feliratozásának szükségességét latolgatják az abban illetékesek. Mégis, érdekes módon, az akusztikai és artikulációs vizsgálatokhoz képest az énekelt magánhangzók észlelésének vizsgálata jóval kevesebb figyelmet kapott eddig a szakirodalomban, és az eddigi eredmények sok kérdést továbbra is megválaszolatlanul hagytak.

Az itt következőkben előbb az éneklés artikulációs és akusztikai sajátosságait vizsgáló kísérleteket vesszük sorra, melyek főként egy-egy nyelv (a svéd vagy az angol) hangzókész-

¹⁰ Mind a beszélőn belüli, mind a beszélők közötti változatosság, annak okai és megvalósulásai számos vizsgálat tárgyát képezték, ráadásul e vizsgálatok száma egyre nő. Ezekből itt csak néhányat említettünk. A beszélőn belüli változatosság egyik aspektusa az életkorral összefüggő fiziológiai és egyéb természetű változások sora, melyek vizsgálata longitudinális (lásd pl. LEE et al. 1999; VORPERIAN–KENT 2007) vagy keresztmetszeti (a magyar vonatkozású vizsgálatok közül lásd pl. BÓNA–IMRE 2010; BÓNA 2012, 2014; GÓSY–BÓNA 2014) lehet. (Míg az előbbi esetben az egyes kísérleti személyek beszédének változását követik nyomon a kutatók, addig az utóbbi esetben az egyes életkori csoportokat nem ugyanazon beszélőkkel reprezentálják.) Egy másik aspektus lehet például az ugyanazon személyek beszédének a beszédstílusok mentén tapasztalható eltérése (lásd pl. MARKÓ et al. 2010). A beszélők közötti változatosság fakadhat például anatómiai, attitudinális, szociális vagy kulturális jellemzőkből, különbségekből is. Ezek vizsgálatára a magyar nyelvvel kapcsolatosan például GÓSY (2004, 2013) munkáit említhetjük.

letét elemezték, több esetben pedig csak egy adott nyelv egyetlen magánhangzóját. Az összefoglalás célja, hogy képet adjon az éneklés specifikus artikulációs jellemzőiről, valamint hogy bemutassa az eddigi eredményeket, melyek a szoprán hangosztályú énekesek akusztikus stratégiáira (elsősorban formánsstratégiáira) vonatkoznak. A bemutatásban előbb a gégeszintű működésekről, valamint a gégeműködés és a szupraglottális artikuláció összefüggéseiről lesz szó, majd az akusztikumra (azaz a formánsfrekvenciák alakulására) vonatkoztatott, közvetett artikulációs és közvetlen akusztikai bizonyítékokat, eredményeket ismertetjük.

Ezután áttérünk az énekelt magánhangzók azonosítását vizsgáló kísérletek eredményeinek ismertetésére, melyek az énekelt magánhangzók sajátos artikulációs és akusztikai tulajdonságainak percepciós vetületét vizsgálják.

Az *1. fejezet* zárásaképp (valamint az eddigi nemzetközi eredmények összefoglalásaként) kiemeljük a máig vitásnak tűnő vagy nyitva maradt kérdéseket, megfogalmazzuk az éneklésnek mint alternatív beszédképzési módnak a magyar magánhangzókra vonatkoztatva relevánsnak tűnő (produkciós és percepciós területeket egyaránt érintő) vizsgálati kérdéseit, és közöljük a jelen értekezés vizsgálatainak alapját képező hipotéziseket.

1.2.1. Az énekelt magánhangzók artikulációs vizsgálata

A következőkben elsőként tehát az énekelt magánhangzók korábbi artikulációs vizsgálatait mutatjuk be. Ebben az alfejezetben elválasztva kezelünk két kérdést, és ennek megfelelően az artikulációs vizsgálatokat is két részre bontva ismertetjük. Egyrészt külön tárgyaljuk azokat a vizsgálatokat, amelyek célja elsősorban nem az akusztikumra, nem az $F_1 : f_0$ formánshangolásra vagy más formánsstratégiákra vonatkozó indirekt bizonyítékok keresése volt (ennek szenteljük a jelen, azaz a *1.1.1. alfejezetet*), hanem egyéb artikulációs stratégiák feltérképezése. Másrészt pedig külön tárgyaljuk azokat a kísérleteket, melyek artikulációs adataikkal elsősorban a szoprán énektechnikában rejlő sajátos rezonanciastratégiákat kívánták – közvetett módon – elemezni (ezeket az *1.1.2. alfejezetben*, az akusztikai sajátosságok alatt tárgyaljuk).

1.2.1.1. A VERTIKÁLIS ÉS HORIZONTÁLIS GÉGEHELYZET ÉNEKLÉSKOR

Már az első tudományos érdeklődésű énektechnikai szakkönyvek említést tesznek az éneklés során alkalmazandó megfelelő függőleges gégehelyzetről. Az énektechnika tudományos vizsgálatának úttörője Manuel GARCIA a 19. században egy, az addigiakhoz képest gyökeresen új, tudományos alapokon nyugvó tanításmódszertan kidolgozását tűzte ki céljául. GARCIA az 1850-es években egy saját maga fejlesztette, kezdetleges laringoszkóp alkalmazásával képes volt megfigyeléseket tenni az éneklés közben tapasztalható fonációra és a függőleges gégehelyzetre, és munkássága nyomán már ekkoriban terjedni kezdett az a nézet, mely szerint

a jó énektechnikával képzett hang produkciója során a gége folyamatosan alacsony helyzetben van (lásd pl. GARCIA 1911). (GARCIA nevéhez fűződik egyébként mindannak a gyakorlata, hagyománya, hogy a gége helyzetére vonatkozó vagy más, ehhez hasonló anatómiai, fiziológiai megfigyelések egyáltalán bekerültek a tanítás-módszertani és énektechnikai szakkönyvekbe, vö. MITCHELL 2005.)

A tudományosabb, állításait anatómiai alapokra is helyező tanítás-módszertani irodalom, valamint az anekdotisztikus szájhagyomány alapján úgy tűnik, hogy a GARCIA által elterjesztett nézet a mai napig tartja magát. Eszerint az énekléskor a magas gégehelyzetben keltett fonáció károsítja a hangszalagokat, így kerülendő – az énekléshez megfelelő gégepozíció a gége lesüllyesztésével érhető el (vö. pl. VENNARD 1967, idézi PABST–SUNDBERG 1992). Azt azonban már régóta tudjuk, hogy a fonáció nem független a szájüregi artikulációtól: mind a gégehelyzetre, mind a hangszalagrezgésre hatással van az artikulációs szervek, az artikulátorokat mozgató izmok tevékenysége (vö. pl. HONDA 1983; HOOLE–KROOS 1998). Kérdés tehát, hogy az énekesek tudományosan is igazolható módon valóban képesek-e állandó (ráadásul alacsony) pozícióban tartani a gégejüket alaphézfrecvenciától és magánhangzó-minőségtől függetlenül; és kérdés az is, hogy valóban célja-e ez az énekesnek, melyre a tanítás során őt trenírozni kell.

A gégehelyzetet az éneklésben objektív módon, modern eszközökkel (például különféle képalkotó eljárásokkal) vizsgáló szakirodalom alapján azt látjuk, hogy a fentebb leírt feltételezés nem bizonyítható maradéktalanul. Mi több, a vizsgálatok eredményei alapján sok esetben éppen a hagyományos feltételezés cáfolata fogalmazható meg, a kutatások ugyanis sok esetben azt igazolták, hogy az éneklésben a gégehelyzet (valamilyen módon) függ az alaphézfrecvenciától.

Az énekelt magánhangzók ejtésekor mérhető vertikális gégehelyzetet (vertical larynx position, VLP) először JOHANNSON és munkatársai (1982) vizsgálták röntgenográfia segítségével. A vizsgálatban a szerzők a magánhangzótér artikulációs és akusztikai értelemben is szélső értékeit képező /a i u/ megvalósítását elemezték egy szoprán és egy alt énekesnél, a szoprán esetében 230 Hz, 470 Hz és 950 Hz alaphézfrecvenciákon, míg az alt esetében 150 Hz, 300 Hz és 600 Hz alaphézfrecvenciákon, laterális röntgenfelvételek segítségével. (Bár az elemzésekkor a gége függőleges pozícióján kívül a kutatók az ajak- és állkapocsnyitás szögét is mérték, valamint a toldalékcső keresztmetszetének kiszámítása után a toldalékcsőre az egyes artikulációs konfigurációkban jellemző rezonanciafrecvenciákat is megbecsülték, a vizsgálat ezen részeivel itt nem foglalkozunk – minderről mint a formánsstratégiákhoz kapcsolódó eredményekről a következő alfejezetben lesz szó.) Az egyetlen szoprán énekes résztvevő eredményei alapján JOHANNSON és munkatársai (1982) megállapították, hogy az elvárásokkal ellentétes módon a gége függőleges helyzete nem volt állandó az alaphézfrecvencia függvényében (ahogy az énektechnikai szakkönyvek alapján feltételezhetnénk), valamint nem igazolódott az sem, hogy a gége a nyugalmi helyzethez képest jellemzően alacsonyabb pozícióban helyezkedett volna el az éneklés során. Az eredmények szerint ugyanis az alaphézfrecvencia emelésével párhuzamosan a gége is megemelkedett, mégpedig az ejtett magánhangzó

minőségétől nagyjából függetlenül, azaz minden esetben. Emellett az egyes magánhangzók esetében némileg eltérő gégepozíciót detektáltak a szerzők: a legalacsonyabb VLP-értékeket az /u/ esetében mérték, míg az /a/ és /i/ esetében ennél valamivel magasabbat. Mivel a kontrollként vizsgált, valamivel alacsonyabb alapfrekvenciákon éneklő alt esetében nem látszott korreláció a vízszintes gégehelyzet és az alapfrekvencia között, a szerzők arra következtettek, hogy a gége emelkedését elsősorban a magasabb alapfrekvenciák idézik elő.

Egy későbbi vizsgálatban PABST és SUNDBERG (1992) egy többcsatornás EGG berendezéssel (tracking multi-channel electroglottograph, TMEGG) vizsgáldott tovább két professzionális szoprán, egy mezzoszoprán és további négy professzionális, valamint egy amatőr bariton énekes résztvevő gégemozgásainak elemzésével. Csak a magasabb alapfrekvenciák és női hangosztályok énekesreire, tehát a szoprán és a mezzoszoprán énekesekre koncentrálnak össze-foglalásunkat a vizsgálat eredményei egyértelműen arra mutattak, hogy a gége függőleges helyzete az alapfrekvencia függvényében nem tekinthető állandónak, és nem igazolódott az a várakozás sem, mely szerint mind a magasabb, mind az alacsonyabb alapfrekvenciákat egyaránt alacsony gégehelyzettel képeznék az énekesek. Ugyanakkor az eredmények alapján a nagymértékű variabilitáson túl nem volt megállapítható semmilyen általános érvényű tendencia sem az alapfrekvencia és a VLP kapcsolatára. Ráadásul variabilitás nem is csak az énekesek között mutatkozott (nem arról volt szó tehát, hogy minden énekes sajátos, de minden kondícióban következetes stratégiát alkalmazott volna), hanem az egyes énekesek produkcióján belül is: a szerzők egymástól gyökeresen eltérő összefüggéseket (vagy az összefüggések teljes hiányát) tapasztaltak az f_0 és a VLP között a kétoktávnyi oktávmenet, az öt-hangnyi skálázás, illetve az oktávugrás énekfeladatokban az egyes adatközlők esetében.

Konklúziójukban a szerzők mindösszesen annyit fogalmaztak meg, hogy az igazán magas alapfrekvenciáknál minden énekesnél látható volt a gége megemelkedése, valamint kifejtették, hogy eredményeik alapján bizonyítottnak vélik azt az állítást, mely szerint az alacsony gégehelyzet egyáltalán nem szükségszerű velejárója a professzionális és jó technikával képzett énekhangnak.

Ráadásúként PABST és SUNDBERG (1992) a hangerő és a gégehelyzet összefüggéseinek vizsgálatára messa di voce énekfeladatokat is rögzítettek (ekkor az énekes egy adott alapfrekvencián fokozatos hangosítást, crescendót, majd halkítást, azaz diminuendót hajt végre), valamint a tüdő vitálkapacitásának maximális kimerítését is kérték az egyes skáláknál, hogy megfigyeléseket tehessenek a gégehelyzet és a tüdőterefogat összefüggéseire is. Az első kérdésben a szerzők azt találták, hogy az intenzitás csak néhány énekes esetében volt hatással a gége függőleges helyzetére, mégpedig úgy, hogy a halkítás során a gége süllyesztését, a hangosítás során pedig a gége megemelését lehetett tapasztalni. A második kérdésben azonban azt a megfigyelést tették, hogy a tüdőterefogat és a függőleges gégehelyzet gyakorlatilag minden énekes esetében szorosan összefüggött: nagyobb tüdőterefogat esetén a VLP alacsonyabbnak mutatkozott, míg a tüdőterefogat csökkenésével (tehát a kilégzéssel) egyértelműen emelkedni látszott. Ezt az összefüggést a szerzők a *tracheal pull* ('légcsövi húzás') néven emlegetett jelenség bizonyítékaként értelmezték, mely szerint a tüdő terfogat-növekedése

(a belégzés) húzó hatást fejt ki a tüdő felfüggesztésére, azaz a gégére, és azon belül is elsősorban a tüdőhöz közvetlenül kapcsolódó gyűrűporcra,¹¹ aminek következtében a tüdőtérfogat befolyásolja a gégehelyzetet, sőt a gége belső mozgásait is. Ha azonban a gégehelyzet és a gége belső mozgásai valóban függenek a tüdőtérfogattól, az nem jelent mást, mint azt, hogy a tüdő térfogata hatással van a zöngképzésre is (vö. SUNDBERG et al. 1988). Amint azt egy bekezdéssel lejjebb látni is fogjuk, ez valóban így van, a szakirodalom erre az összefüggésre ugyanis számos közvetlen bizonyítékkal is szolgál.

HURME és SONNINEN (1995) röntgenképek segítségével vizsgálta mind a gége vertikális (VLP), mind pedig horizontális (szagittális) elmozdulását (sagittal larynx position, SLP) az alapfrekvencia függvényében tizenkét énekesnél, akik közül kilenc tartozott a szoprán vagy mezzoszoprán hangosztályba. Az énekesek a felvételek során az /a/ magánhangzót ejtették emelkedő skálák hangjain. A korábbiakhoz hasonlóan ez a vizsgálat is igen nagy variabilitást mutatott ki az egyes énekesek között. Míg két szoprán esetében a mozgások összessége egy, az alapfrekvencia emelésével hátra (posterior irányba) és felfelé (superior irányba) irányuló elmozdulást mutatott, négy esetben inkább előre (anterior irányba) és felfelé (superior irányba) mozdult az énekesek gégéje. Mi több, egy esetben a mozgások összessége inkább hátra (posterior irányba) és lefelé (inferior irányba) mutatott, míg egy további esetben egy, leginkább rendszertelennek tekinthető, komplex cikkek mozgást tapasztaltak a kutatók. HURME és SONNINEN (1995) eredményei alapján tehát – a korábbiaknak megfelelően – a konstans alacsony gégehelyzetre vonatkozó hagyományos feltételezés nem nyert megerősítést, valamint – ismét csak a korábbi eredményekkel egybecsengő módon – a szerzők az f_0 és a VLP, illetve az f_0 és az SLP összefüggéseinek számos lehetséges megvalósulását demonstrálták. (Mindeközben ráadásul a vizsgálat arra is rávilágított, hogy a függőleges gégehelyzet önmagában nem elegendő a „gégehelyzet” leírásához.) A szerzők végső konklúziója szerint a gégehelyzet és a magas alapfrekvenciájú éneklés összefüggéseiben a lényegi pont feltehetőleg nem a gégemozgások konkrét útvonala (és összefüggése az f_0 -lal vagy az intenzitással), hanem az egymásra ható, az izmok által kifejtett erők kiegyensúlyozása.

Amint arra az előbbieken már több helyütt is utaltunk, a függőleges gégehelyzet több okból is meghatározó lehet a produkciós folyamatban, ezért is találjuk a szakirodalomban azt, hogy a gége helyzete az éneklésben egy fontosnak látszó, sokat vizsgált kérdés.

¹¹ Valamivel részletesebben kifejtve a *tracheal pull* terminus a következőkre utal: mivel belégzéskor a rekeszizom megfeszül, és lefelé húzza a tüdőlebeenyeket, a tüdő is húzó hatást fejt ki a felfüggesztésére, azaz a gégére, mégpedig úgy, hogy lefelé húzza azt, és különösen a gégeporcokból a hozzá legközvetlenebbül kapcsolódó gyűrűporcot, így egyúttal távolítja is a gyűrűporcot a pajzsportól. Ez a húzó hatás tehát az alapfrekvencia emeléséhez, illetve a magas alapfrekvencia megtartásához szükséges mozgásoknak, azaz a pajzsporc előrebillentésének, valamint a pajzsporc a gyűrűporchoz közelítésének ellene hat, megnehezíti azt. E hatás bizonyítása érdekében SUNDBERG és munkatársai (1988) a cricothyroideus izom működését regisztrálták a légzéssel összefüggésben kitartott (állandó alapfrekvenciájú) hangok ejtésében. A szerzők azt találták, hogy közvetlenül a belégzés után a cricothyroideus izom jobban megfeszül, ami arra utal, hogy az énekesnek/beszélőnek a gyűrű- és pajzsporc közti állandó távolság fenntartásához, azaz az alapfrekvencia-csökkenés elkerüléséhez a belégzést követően egy, a gyűrűporcot érő lefelé húzó hatást kell kompenzálnia (SUNDBERG et al. 1988).

A gége helyzete által meghatározott egyik fontos produkciós tényező – amint arra fentebb is tettünk már utalást – a fonáció. Míg vizsgálataiban ZENKER és ZENKER (1960), valamint PABST és SUNDBERG (1992) azt demonstrálta, hogy a lesüllyesztett gégehelyzet a hangszalagok nagyobb fokú távolodását (abdukcióját) eredményezi, SUNDBERG és ASKENFELD (1981) arról számolt be, hogy a magas gégehelyzet a hangszalagok nagyobb mértékű közelítéséhez (addukciójához), azaz feszítettebb, préseltébb fonációhoz vezet. (Éppen ezért egyébként az a feltételezés, mely szerint a magasabb gégehelyzet az éneklésben kerülendő, első közelítésben valóban helyesnek látszik.) Mindebből az következik, hogy ha a gége helyzete (például a levegővel telt tüdő által kifejtett húzó hatás következtében) függőleges irányban változik, az mind a fonáció módjára, mind pedig a létrejövő hang hangszínzetére nagy hatással van.

A gége helyzete által meghatározott másik fontos tényező a toldalékcsovi rezonanciák frekvenciaértéke. Mivel a gége emelése rövidíti, süllyesztése pedig nyújtja a toldalékcsovet (a garatüreg méretének befolyásolásán keresztül), így a gége föl-le mozgatása hatással van a formánsfrekvenciák értékére is: a gége süllyesztése csökkenti, emelése pedig növeli a toldalékcso sajátrezonanciáinak értékét. A kísérleti eredmények szerint, melyek erről az általános összefüggésről pontosítják az információinkat, a gége emelése növeli a zárt vagy első nyíltsági fokú palatális hangzók F_2 -értékét, növeli mind az F_1 , mind pedig az F_2 értékét a középnyílt és nyílt, vagyis a harmadik és negyedik nyíltsági fokú magánhangzók esetében, valamint több hangzó esetében is az F_3 és F_4 értékének egyidejű növekedését okozza (SUNDBERG–NORDSTRÖM 1976). Ezzel összefüggésben az is valószínűsíthető, hogy a vertikális gégehelyzet játszik szerepet a férfi hangkategóriák énekprodukciójára jellemzőnek tartott énekesformáns kialakításában is.

Az énekesformáns a nevével ellentétben valójában nem önálló rezonanciafrekvencia, hanem formánsklaszter (nagy sávzélességű, nagy intenzitású frekvenciasáv), mely az F_3 , F_4 és F_5 formánsokból jön létre. A klaszter létrejöttének oka, hogy a mindennapi beszéd során egyébként nem változtatott (mert a szájüregi artikulátorokkal nem vagy alig befolyásolható) frekvenciaértékű F_3 , F_4 és F_5 formánsok a mélyebb alaphangú (férfi hangosztályokhoz kötődő) éneklésben frekvenciájukban közelednek egymáshoz, és ennek eredményeképpen (pusztán a csatolt rezgések fizikai törvényszerűségeinek következtében) egyúttal egymás amplitúdóját is megemelik. A csatolás következtében tehát az F_3 , F_4 , F_5 formánsok helyek frekvenciasávja többlet energiát kap (ez megközelítőleg a 2-2,5 és 3-3,5 kHz közötti tartomány), így az énekesformánssal bíró énekhang felerősödik, mégpedig éppen abban a frekvenciatartományban, melyben a kísérő zenekar már nem képes maszkolni azt (hiszen az énekesformáns frekvenciasávján a hangszerek hangja már nem tartalmaz intenzív összetevőket). Az énekesformánsnak köszönhetően tehát a mélyebb hangosztályok énekeseinek hangja akár nagyzenekari kíséret mellett is jól hallható lesz (SUNDBERG 1972, 2001). Amint azt a magánhangzók képzésekor már tárgyaltuk, a szájüregi artikulátorok alapvetően csak az F_1 és F_2 formánsok frekvenciaértékét módosítják, az F_3 , F_4 és F_5 formánsok értékének nagyobb mértékű befolyásolására nem képesek – így nem feltételezhetjük azt sem, hogy az énekesformáns kialakításáért ezek az artikulációs szervek lennének a felelősek. A függőleges gégehelyzet azonban

a magasabb sorszámú rezonanciáknak a frekvenciaértékét is képes módosítani (lásd SUNDBERG–NORDSTRÖM 1976), éppen ezért a szakirodalomban meglehetősen régóta élő, elterjedt feltételezés az, hogy az énekesformáns kialakítását a függőleges gégehelyzet változtatásával érik el az énekesek, mégpedig olyan módon, hogy azzal a garatüreg és a gégecső arányát megközelítőleg 6 : 1 arányra állítják (SUNDBERG 1975, 2001; TITZE 1998, 2001).

Az énekesformáns meglétét ugyanezen szakirodalom csak a férfi hangosztályok énekesei-nek produkciójában véli igazoltnak. SUNDBERG (2001) szerint a női hangosztályok, elsősorban a szopránok hangja e kérdés tekintetében egyáltalán nem mutat egységes képet. SUNDBERG összefoglalásában azt írja, hogy bár némely vizsgálat sikerrel mutatott ki valamifajta intenzitáserősödést a magasabb frekvenciakomponensekben a szopránok esetében is, ennek az erősödésének a mértéke korántsem volt a férfi hangosztályokban tapasztaltakhoz mérhető. Mindemellett a szerző szerint feltételezhető, hogy ez az intenzitáserősödés hozzájárul a női énekhang hallhatóságának növeléséhez nagy terekben és nagyzenekari kíséret mellett (vö. SUNDBERG 1977, 2001; THORPE et al. 2001; WEISS et al. 2001). A szoprán hangokban megjelenő énekesformáns kérdéséhez ráadásul még egy további alapvetően fontos megjegyzés is hozzáfűzhető. Nem szabad megfélemednünk ugyanis arról sem, hogy az énekesformáns vizsgálata a szoprán hangokban eleve sokkal nehezebb feladat, mint a férfi hangokban, egyszerre két okból is. Egyrészt azért, mert a spektrális alul-mintavételezés miatt a felsőbb formánsok detekciója a magas alapfrekvenciájú női éneklésben eleve akadályozott, másrészt pedig azért, mert az $F_1 : f_0$ hangolás a magasabb női énekhangok spektrumában már eleve erősödést okoz, amitől a magasabb sorszámú rezonanciákból kialakított énekesformáns feltételezett hatása gyakorlatilag disszociálhatatlan. Ilyen értelemben tehát valójában feltehetőleg nincs is relevanciája az énekesformáns meglétét firtatni a női hangosztályok esetében, hiszen valószínűsíthető, hogy nem ez az eszköz segíti hozzá a női szolistákat a zenekaron is átható, hangos énekprodukciónak, hanem sokkal inkább a már említett $F_1 : f_0$ hangolás. Éppen ezért a jelen értekezésben, azaz a szoprán énekesek énekprodukciónak vizsgálatában nem foglalkozunk az énekesformáns megjelenésének kérdésével.

1.2.1.2. A GÉGE BELSŐ MOZGÁSAI, A NYELVCSONT SZEREPE AZ ÉNEKLÉSBEN

HURME és SONNINEN (1995) a gége helyzetével párhuzamosan a gégeporcok egymáshoz viszonyított helyzetét, valamint a nyelvcsont helyzetét is vizsgálták, tehát megfigyeléseket tettek a gége egyes belső mozgásaira is. Míg a nyelvcsont szerepe a beszédben és az ember élettani működései szempontjából a mai napig némiképp rejtélyes (vö. SEIKEL et al. 2010), addig HURME és SONNINEN (1995) az éneklésben kapott adatai – kiegészítve HONDA a témában folytatott vizsgálataival (vö. pl. HONDA 1983, lásd alább) – e csontocska egy (vagy több) lehetséges funkciójára is utalni látszanak.

HURME és SONNINEN (1995) azt találta, hogy a magasabb alapfrekvenciákat nagy arányban magasabb gégehellyel produkáló énekeseknél a nyelvcsont a magas alapfrekvenciákon

a beszédben nem tapasztalható mértékben a pajzsporc elé billent, ráadásul olyannyira, hogy annak elülső, legalacsonyabban elhelyezkedő pontja akár a pajzsporc csúcsát (azaz az ádámcsutkát) is elérhette. Bár a szerzők a nyelvcsont mozgását az f_0 -lal szoros összefüggésben figyelték meg, mégis úgy vélték, hogy ezeket a mozgásokat elsősorban nem az f_0 -kontroll, hanem a VLP változása befolyásolta. Következtetésük szerint ugyanis a nyelvcsont előrebillentésének célja magas gégehelyzetben az akadálytalan levegőáramlás biztosítása a gégecső tágítása révén. (Következtetésüket a szerzők arra alapozták, hogy a levegő áramoltatása megemelt gégehelyzetben a gége összepréselődése miatt nehezebb.)

Egy másik elképzelés szerint azonban a nyelvcsont az f_0 befolyásolásával van szoros összefüggésben. HONDA (1983, é. n.) vizsgálatai szerint ugyanis az f_0 extrém emelésekor (amire például énekléskor van szükség) a nyelvcsonthoz kapcsolódó izmok (többek közt a geniohyoideus vagy a genioglossus) is megfeszülnek, és előre húzzák a nyelvcsontot, azért, hogy az egyébként normál körülmények között (de kisebb hatásfokkal) az alaphfrekvenciát emelő cricothyroideus izom működését, a pajzsporc előrebillentését fokozzák. Ha ugyanis a nyelvcsont előrebillen, és lefelé húzódik, a nyelvcsont és a pajzsporc között lévő rugalmas szövetek (a thyrohioideus izom, a középső és szélső thyrohioideus szalagok, a thyrohioideus membrán stb., vö. 2. ábra) segítségével a nyelvcsont magával húzza a pajzsporcot is. Így a pajzsporc a normál rotációs fokánál nagyobb szögben képes elmozdulni, miközben egyúttal jobban megfeszíti a hangszalagokat is, tehát a nyelvcsont előrebillentésével a „szokásosnál” magasabb alaphfrekvenciák érhetőek el (HONDA 1983, é. n.). HONDA szerint a nyelvcsont helyzetének irányítása egy igen komplex rendszerben zajlik, ahhoz ugyanis nem csak a nyelvcsontot előrébb mozdító geniohyoideus vagy genioglossus izmok megfeszítése szükséges, hanem az ezen izmok antagonistá (ellentartó) izmait képező, a nyelvcsontot hátra, a garat felé húzó középső garatösszeszorító izom (constrictor pharyngis medius) és a nyelvcsontot lefelé húzó sternohyoideus működtetése is (HONDA é. n.).

A gége, valamint a nyelvcsont helyzetét és funkcióját vizsgáló kutatások egyrészt hozzájárulnak ahhoz, hogy jobban megérthessük az éneklésnek mint különleges produkciós módnak néhány magyarázatra szoruló aspektusát, másrészt pedig rávilágítanak egy, a beszéd- és énekprodukción vizsgálatait alapvetően meghatározó összefüggésre is.

Ezek a vizsgálatok részben választ adnak arra a talán alapvetőnek tűnő kérdésre, hogy hogyan képesek a szoprán énekesek az énekmódszertanilag nem képzett beszélők által produkálhatónál magasabb alaphangok kiadására és az alaphfrekvencia-emelésnek a hétköznapi-nál nagyobb fokú kontrolljára. Az e kérdésre adható válasz egyik része ugyanis az lehet, hogy az énekesek nem csak a normál beszédmódban (és egyébként minden beszélő által) használt cricothyroideus izom működtetésével emelik és kontrollálják hangjuk alaphfrekvenciáját, hanem a nyelvcsont összeköttetéseit kihasználva a fent leírt komplex rendszer egészének begyakorlott manipulálásával is.

Ugyanezen vizsgálatok a gége helyzetének és működésének elemzésén keresztül azonban arra a fontos tényre is rávilágítanak, hogy hiába elegendő bizonyos kérdések vizsgálatahoz és megválaszolásához a beszédképzés lineáris, a forrást és a szűrőt, tehát a géget és

a szupraglottális artikulációs rendszert élesen elválasztó modellje (melyben a gége működésére elvben nem hat a szupraglottális rendszer tevékenysége), ez a leképezés sokszor túlságosan is leegyszerűsítő. Nem szabad ugyanis megfedkezünk arról, hogy a teljes artikulációs rendszer működése a valóságban ilyen értelemben nem lineáris, hiszen a rendszer egyes részei egymás tevékenységét is (közvetlenül vagy közvetetten) befolyásolva vesznek részt az emberi hangadás folyamatában.

1.2.2. Az énekelt magánhangzók akusztikai vizsgálata: rezonanciák a szopránénekülésben

A magas alaphfrekvencián énekelt magánhangzók akusztikai megvalósulásának esetében, amint már említettük, a fő kérdés az, hogy mi történik akkor, ha az f_0 magasabbra emelkedik, mint a beszédben jellemző F_1 értéke ($F_{1[\text{beszéd}]}$). Mivel a szoprán énekesekre írt művek hangterjedelme átlagosan a c' (262 Hz) zenei hangtól a c''' (1047 Hz) zenei hangig terjed, a szopránok énekprodukcója során ez a kérdéses helyzet gyakran előáll.¹² Tudjuk azt, hogy ha az F_1 az f_0 értéke alá kerülne, tehát az énekes nem változtatna a magánhangzónak a beszédben megszokott ejtésén a megemelt alaphfrekvencián, az énekesnek három nagyon fontos, az énekelt hangzók nyelvi és esztétikai minőségét is befolyásoló következménnyel kellene számolnia. Az egyik következmény egy, a magánhangzó-minőség kialakítása és észlelete szempontjából igen fontosnak tartott (lásd pl. GÓSY 2004) akusztikai kulcs, az F_1 elvesztése, a másik az akusztikai energia, a hangnyomásszint, tehát a hangosságérzet csökkenése, a harmadik pedig a hangszínezetben hirtelen beálló változás lenne (vö. SUNDBERG 1979, 1987). Ennek megfelelően a szakirodalom a szopránok magánhangzóejtéséről azt feltételezi, hogy (magas alaphangokon) az alaphfrekvencia függvényében változik, mégpedig úgy, hogy ha az f_0 értéke magasabb lenne, mint az $F_{1[\text{beszéd}]}$, a szoprán énekes az F_1 értékét felhangolja a megemelt f_0 értékére vagy valamivel a fölé (lásd pl. SUNDBERG 1987).

Ahogy az már korábban említettük, a kérdés vizsgálatának úttörője, SUNDBERG (1979, 1987) szerint annak kedvező akusztikai következményei miatt az $F_1 : f_0$ hangolásról szóló feltételezés empirikus bizonyítékok nélkül is valószínűsíthető, és szerinte az is könnyen belátható, hogy az énekesek nagyon sokat profitálhatnak abból, ha énektechnikájukat ennek megfelelően, azaz az üregi rezonanciák jobb kihasználásával fejlesztik. Amint azonban arra már szintén utaltunk, a felvetés alátámasztása empirikus bizonyítékokkal meglehetősen nehéz, mert a magas alaphfrekvenciákon ejtett magánhangzók spektrálisan alul-mintavételezettek, és ezért azokban a formánsfrekvenciák középpértékének meghatározása a szájnnyílást elhagyó akusztikus jel hagyományos, Fourier-transzformáción alapuló elemzésével nem vezet megbízható eredményre (vö. DE CHEVEIGNÈ–KAWAHARA 1999, valamint a jelen értekezés *1.1. alfejezete*).

¹² A zárt képzésű magyar magánhangzók első formánisa a toldalékcso hosszától, a beszédstílustól és beszédtempótól is függően nagy variabilitást mutathat, de átlagosan 200 és 500 Hz közötti értéken realizálódik (lásd pl. GRÁCFI–HORVÁTH 2010 és a jelen könyv 2., illetve 6. fejezete).

A fenti okokból az egyes nyelvek magánhangzóinak minőségét, azaz akusztikai értelemben az első két (vagy három) formánsát a magas alapfrekvenciájú éneklés során érintő változásokról szóló információink forrása kettős. A vizsgálatok egyik (és talán valamivel nagyobb) része inkább közvetettnek tekinthető információkkal szolgált, mert ezek elsősorban az énekelt magánhangzók artikulációját vizsgálták, nem pedig az így előálló akusztikus szűrő tulajdonságait (bár néhány esetben az artikulációs adatokból a szerzők fizikai és matematikai modellek segítségével a magánhangzók valószínűsíthető formánsfrekvenciáit is igyekeztek megállapítani). A vizsgálatok másik csoportja ezzel szemben elsősorban akusztikai elemzést végzett, és ezek révén igyekezett meghatározni az énekelt magánhangzók formánsértékeit – ilyen értelemben tehát ezek inkább a magánhangzók akusztikai szerkezetét kialakító akusztikus szűrőről közvetlen információkat szolgáltatató elemzéseknek tekinthetők. Mivel a magas alapfrekvenciájú beszédjelet spektrális alul-mintavételezettség jellemzi, ezért ez utóbbi vizsgálatok a hagyományos, a kisugárzott jel Fourier-transzformációján alapuló akusztikai elemzéstől eltérő, újszerű vagy nem szokványos méréseket jelentettek. A következőkben mindezeket, a magas alapfrekvencián énekelt magánhangzókat rendre direkt és indirekt módszerekkel vizsgáló kísérleteket és a vizsgálatok eredményeit tekintjük át. (Azokban az esetekben, ahol egy vizsgálaton belül a kutatók mind artikulációs, mind pedig valamely nem szokványos akusztikai mérés segítségével nyert akusztikai adatokra nézve is tettek megfigyeléseket, a kétféle módon nyert eredményeket szétválasztva közöljük, azt szem előtt tartva, hogy az egyes eredmények inkább direkt vagy inkább indirekt bizonyítékokként értelmezhetők a toldalékcso rezonanciáinak alakulására nézvést.)

A szoprán énekesek a formánshangolást előidéző artikulációs stratégiáival kapcsolatban SUNDBERG (1975), JOHANSSON és munkatársai (1982), SUNDBERG és SKOOG (1995), valamint BRESCH és NARAYANAN (2010) folytattak vizsgálatokat (rendre) egy különleges, egyéni készítésű, a metszőfogak és a két állkapocs távolságát mérő mérőeszköz (vö. COOKER 1973), röntgenográfia, egy magnetométer alapú mérőeszköz (vö. BRANDERUD et al. 1994), valamint mágneses képalkotó eljárás (MRI) segítségével.

SUNDBERG (1975) a svéd /u o e i y/ magánhangzókat vizsgálta 262 Hz, 394 Hz, 523 Hz és 698 Hz alapfrekvencián (a c', g', c'', g'' zenei hangokon) énekelve, egy szoprán ejtésében. Eredményei szerint, míg az /a/ ejtésékor az f_0 emelését közel állandó mértékű állkapocsnyitás kísérte, addig a többi vizsgált magánhangzó esetében az f_0 emelésével párhuzamosan a nyitásszög növekedése volt tapasztalható. Az /u o e i y/ magánhangzókra kapott artikulációs adatokból SUNDBERG arra következtetett, hogy a kísérletben részt vevő énekes az énekelt magánhangzóinak első formánsát az alapfrekvencia és a magánhangzó-minőség függvényében egyre feljebb és feljebb hangolta az f_0 emelésével, ha az f_0 egyébként elérte vagy megközelítette volna az F_1 értékét. Erre a következtetésre a szerzőnek alapot az a mára már közismert összefüggés szolgáltatott, mely szerint az állkapocsnyitás szögének növelésével az első formáns frekvenciaértéke is növekszik (vö. pl. LINDBLOM–SUNDBERG 1971). Az /a/-ra talált kivételes tendenciával kapcsolatban pedig SUNDBERG azt az állítást fogalmazta meg, hogy az az $F_1 : f_0$ hangolás feltételezése mellett is jól magyarázható eredmény. A szerző ennek

megvilágításához a következőképpen érvelt. Mivel az /a/ nyílt (negyedik nyíltsági fokú, legalsó nyelvvállású) hangzó, és ennek megfelelően a beszédben is meglehetősen magas F_1 értékkel realizálódik, ezért nyilvánvalóan az éneklésben is csak magas alapfrekvenciákon esik az $F_1 : f_0$ hangolás hatálya alá. Ha tehát az /a/ esetében az állkapocsnak az f_0 emelésétől függetlenül, nagyjából állandó nyitásszögét találjuk, akkor az nem az $F_1 : f_0$ hangolás cáfolataként értelmezendő, hanem csak azt bizonyítja, hogy a kísérletben vizsgált alapfrekvenciák nem voltak magasabbak az /a/ beszédben jellemző első formánsának értékénél.

JOHANNSSON és munkatársai (1982) a magánhangzótér artikulációs és akusztikai értelemben is szélső értékeit képező /a i u/ megvalósítását vizsgálták egy szoprán és egy alt énekesnél, a szoprán esetében 230 Hz, 470 Hz és 950 Hz alapfrekvenciákon, míg az alt esetében 150 Hz, 300 Hz és 600 Hz alapfrekvenciákon. A vizsgálat során a kutatók az ajak- és állkapocsnitítás szögét, valamint a gége függőleges helyzetét mérték laterális röntgenfelvételek segítségével, valamint ezekből az artikulációs adatokból akusztikai adatokat is származtattak a következőképpen. A szerzők a toldalékcső középvonalának meghatározása után kiszámolták a toldalékcső adott pontjain jellemző oldalnézeti keresztmetszeteket, azaz a toldalékcső keresztmetszefüggvényét (*aperture function*, melyen az egyes szakaszoknak a gégetől való távolságára vetítve a toldalékcső adott pontjain jellemző keresztmetszeteket vesszük fel). Ezek után pedig a kapott keresztmetszefüggvényt keresztmetszeterület-függvénné (*cross-sectional area function*) konvertálták, amihez a hiányzó adatokat, azaz a röntgenfelvétel síkjára merőleges távolságokat a garatüregre és a szájüregre jellemző, korábban meghatározott állandókkal pótolták. Végül a kutatók a toldalékcső keresztmetszeterület-függvénye alapján plexilapokból fizikai csőmodellt készítettek, melyet a „gége felőli” oldalán szinuszhullámokkal gerjesztettek, így a cső másik végén kisugárzott akusztikai jel elemzésével meg tudták határozni a cső sajátfrekvenciáit, azaz a csőmodell alapját képező toldalékcső-konfigurációra jellemző formánsértékeket.

JOHANNSSON és munkatársainak (1982) az F_1 frekvenciaértékének manipulációja szempontjából fontos artikulációs eredményei szerint a szoprán résztvevő következetesen növelte az állkapocsnitítás szögét az f_0 emelésével, valamint mind a nyelvkontúr, mind az ajkak távolsága, mind pedig az ajkak „behúzása” tekintetében uniformizálta az /a/, /i/ és /u/ hangzóit a legmagasabb, 950 Hz alapfrekvencián. Az akusztikai eredmények azonban azt mutatták, hogy ugyananezen énekes nem az f_0 értékére, hanem valamivel az f_0 értéke alá hangolta az F_1 -rezonanciát akkor, amikor az f_0 egyébként magasabb értéket vett volna fel, mint az F_1 beszédben jellemző frekvenciaértéke, ami – ha valóban igaz –, azt jelenti, hogy ezeken az alapfrekvenciákon az F_1 nem jelent meg az énekes magánhangzóiban az akusztikus kimenetben, és valószínűleg az f_0 -t sem erősítette. Ez utóbbi eredményt a szerzők a modelltől fakadó hibaként értelmezték, és értelmezésük szerint az artikulációs adatok egyértelműen bizonyították az F_1 emelésére, valamint az $F_1 : f_0$ hangolásra irányuló produkciós tendenciákat. Mindezen felül a szerzők arról is beszámoltak, hogy a szoprán énekes (szintén az f_0 emelésével) csökkentette az elől képzett, és növelte a hátul képzett hangzók F_2 -értékét is, így

uniformizálva a három beszédhangot a kísérletben vizsgált legmagasabb alapfrekvencián, azaz 950 Hz-en.

SUNDBERG és SKOOG (1995) a svéd /a o u i e/ magánhangzókat vizsgálta egy szoprán és három mezzoszoprán (továbbá egy mezzoszoprán-alt, két tenor, két bariton és egy basszus énekes) ejtésében, két oktáv hangterjedelemben. A kísérletben lefedett hangartomány a szoprán énekesnő esetében a 220 és 880 Hz (azaz az *a* és *a''* zenei hangok) közötti, a mezzoszopránok esetében pedig a 196 és 698 Hz (azaz a *g* és *f''*) közötti alapfrekvencia-tartományt jelentette. A két kutató vizsgálatában a szopránokat és mezzoszopránokat érintő eredmények tekintetében a korábbiakkal csak részben egyező eredményekre jutott, ugyanis az egyes magánhangzókra (valamint az egyes hangosztályok képviselőinek stratégiáira) eltérő tendenciákat mutattak ki a következők szerint. Úgy találták, hogy a magasabb beszédbeli F_1 -gyel realizálódó /a o/ hangoknál egyértelműen tapasztalható volt az állkapocsnyitás szögének növelése akkor, amikor az f_0 megközelítette az F_1 értékét. A beszédben valamivel alacsonyabb F_1 -gyel realizálódó /o e/ esetében azonban ez a tendencia csak magasabb alapfrekvencián, mégpedig körülbelül 5 félhanggal az után kezdődött, hogy az f_0 elérte volna az F_1 értékét az f_0 emelése során. A beszédben a legalacsonyabb F_1 -gyel realizálódó /u i/ esetében találták a szerzők a legkomplexebb és legdivergensebb eredményeket. Az /u/ és /i/ hangok esetében ugyanis azt tapasztalták, hogy minden énekes, köztük a szoprán és mezzoszoprán énekesek is csak az általuk a vizsgálatban lefedett frekvenciatartomány legfelső részében mutattak az f_0 -tól függő, az állkapocsnyitásszög növelésére irányuló produktív tendenciát. A mezzoszopránok esetében az állkapocsnyitásszög növelését célzó tendencia az F_1 beszédbeli értéke fölött 7 félhanggal, míg a szoprán énekes esetében az F_1 frekvenciaértékéhez közel (körülbelül az F_1 frekvenciaértékén) kezdődött az f_0 emelése során. (Érdekességképpen megemlítjük, hogy ugyanezen vizsgálatban az /u i/ hangok esetében a tenorokra a szopránokhoz hasonló tendenciát találtak a szerzők, viszont a bariton és basszus hangosztályú énekeseknél azt látták, hogy az állkapocsnyitásszög növelése már a beszédben jellemző F_1 -érték alatt elkezdődött.)

SUNDBERG és SKOOG (1995) tanulmányában végezetül két következtetést fogalmazott meg. Az egyik szerint, mivel az állkapocsnyitás nemcsak az F_1 értékének módosítására van hatással, hanem a fonációra is, valójában nem meglepő az, hogy nem minden magánhangzó és nem minden énekes hangosztály esetében figyelhető meg az állkapocsnyitás az f_0 értékétől függő növelésének uniform tendenciája. A tanulmány másik konklúziója pedig valójában egy, a jövőben tesztelendő feltételezés megfogalmazása: a szerzők szerint ugyanis a beszédben alacsonyabb F_1 -gyel realizálódó (zárt, első nyíltsági fokú vagy felső nyelvvállású) magánhangzók esetében valószínűsíthető, hogy azok első formánsának értékét nem az állkapocs, hanem más artikulátorok, mégpedig a nyelv manipulációja révén módosítják az énekesek.¹³ Erre

¹³ Ezt a feltételezést a szerzők többek közt LINDBLOM és SUNDBERG (1971) eredményeire alapozzák, melyek szerint noha az F_1 értékének befolyásolására minden magánhangzó esetében a nyelv függőleges helyzetének változtatása volna a leggazdaságosabb artikulációs megoldás, a nyelv süllyesztése csak a zártabb magánhangzók esetében eredményezi egyértelműen az F_1 értékének növelését, a nyíltabb hangzók esetében éppen hogy csökkenti

építve a szerzők arra a következtetésre jutnak, hogy míg az állkapocsnyitás szögének az f_0 emelését kísérő fokozatos növekedése a nyíltabb hangzók esetében az F_1 emelésének közvetett bizonyítékaként értelmezhető, az állkapocsnyitásszög változásának hiánya a zártabb /u i/ esetében nem az F_1 -hangolás hiányára utal, pusztán arra a tényre, hogy a formánshangolást előidéző artikulációs stratégiát az adott módszerrel (az állkapocshelyzet regisztrálásával) nem sikerült közvetve bizonyítani. Így bár SUNDBERG és SKOOG (1995) artikulációs eredményei még közvetlenül sem igazolják a szopránoknak az F_1 hangolására tett erőfeszítéseit az /u/ és /i/ magánhangzók esetében, valamint kétségeket hagynak a hangolás megjelenését illetően az /o/-ra és /e/-re nézvést is, mégis a talált tendenciákat a kutatók általánosan az f_0 emelését kísérő $F_1 : f_0$ hangolás közvetett bizonyítékaiként értelmezik.

BRESCH és NARAYANAN (2010) öt szoprán énekes énekprodukciónak vizsgálta két oktáv hangerjedelemben (233 Hz és 932 Hz, azaz az $a\#$ és az $a\#\#$ zenei hangok között), az /a e i o u/ hangokon skálázva valós idejű mágneses rezonancián alapuló képalkotó eljárás (MRI) segítségével, vibrató nélküli ének-módban. BRESCH és NARAYANAN az MRI segítségével készített szagittális felvételekből kiindulva az énekelt magánhangzók artikulációját, valamint akusztikai megvalósulását is vizsgálta. A szerzők az akusztikai mérésekhez az oldalnézeti képek segítségével JOHANNSON és munkatársainak (1982) megoldásához hasonlóan megalkották a toldalékcso akusztikus csőmodelljét, ám ezt nem modellezték a fizikai valóságban, pusztán matematikai úton, és ebből a matematikai modellből számították ki a csőhöz tartozó formánsfrekvenciákat.

Artikulációs adataik alapján BRESCH és NARAYANAN (2010) megállapították, hogy az f_0 emelésével minden énekes egyre nyitottabb és egymástól egyre kevésbé eltérő artikulációs beállítással ejtette az öt különböző minőségű magánhangzót. A modell alapján számított akusztikai adataikat (egészen pontosan az F_1 értékét) az f_0 értékével összevetve a szerzők azt találták, hogy csak az /i/ és /u/ F_1 -e esetében lehetett minden énekes esetében szignifikáns korrelációt látni. Az /e/ és /o/ esetében már csak két énekesre kaptak szignifikáns összefüggéseket, míg az /a/ esetében egyetlen adatközlő F_1 - és f_0 -értékeire sem.

BRESCH és NARAYANAN eredményeikből arra következtettek, hogy csak az /i/ és /u/ esetében mondható ki egyértelműen az F_1 -nek az alapfrekvenciától függő általános (tehát nem énektechnika- vagy énekesfüggő) hangolási stratégiája. A kutatás eredményeinek értelmezéséhez azonban még két tényezőt fontos látnunk. Az egyik – és ezt részben a szerzők is hangsúlyozzák –, hogy az /a/ esetében a korreláció, azaz a hangolási tendencia hiánya a várt eredmény. Amint a korábbi vizsgálatok is egybehangzóan tanúsították, az /a/ és az /a/ ejtése stabilabb, azaz az alapfrekvenciától nagyobb mértékben független, mert ezek nyitottabb ejtésű, magas első formánssal rendelkező hangzók, tehát ezek F_1 formánsát csak magas alapfrekvenciák éneklésekor éri el az f_0 akkor, ha az énekes emeli énekhangjának alaphangját. A másik tényező pedig az, hogy nyíltsági fokukból kifolyólag valójában az /e/ és /o/ esetében is csak

azt. Éppen ezért a szerzők azt feltételezik, hogy a nyíltabb hangzóknál az F_1 emeléséhez inkább az állkapocsnyitásszögét kell növelni, míg a zártabbaknál ugyanezen akusztikai célra a nyelvhat süllyesztése a leginkább valószínűsíthető artikulációs stratégia.

a magasabb, 4-500 Hz fölötti alapfrekvenciákon várnánk az F_1 -nek a beszédbeli átlagos frekvenciájánál magasabb frekvenciájú megjelenését, hiszen az /e/ és /o/ hangok F_1 -értéke is átlagosan 4-500 Hz körül jelenik meg. Ez BRESCH és NARAYANAN vizsgálatára nézve azt jelenti, hogy az ott vizsgált alapfrekvenciák alsó tartományában, azaz kvázi az első oktávon és még valamivel a fölött sem várnánk hangolást az /e/ és /o/ esetében sem, így ezen a tartományon korrelációt sem jósolhatunk az F_1 és f_0 értékek között. Következésképpen feltételezhetjük, hogy a 16. a korrelációs számításnál vizsgált adatpont (azaz a 16 alapfrekvencia-érték és az ahhoz tartozó F_1 -érték) második felének összefüggéseit, azaz a második, magasabb oktávban talált hangolási tendenciákat BRESCH és NARAYANAN vizsgálatában elfedhették a hangolás szempontjából nem kiritikus tartományon (azaz az első oktávon) talált tendenciák. Ilyen módon tehát BRESCH és NARAYANAN (2010) eredményei valójában nem az $F_1 : f_0$ formánshangolási stratégia hiányát bizonyították az egyes hangzók esetében, hanem csak azt az elképzelést erősítették meg, mely szerint az $F_1 : f_0$ hangolás megjelenése az F_1 frekvenciaértékétől, tehát a hangok nyíltsági fokától függ.

BRESCH és NARAYANAN következtetések között azt is megjegyezték, hogy a vizsgálatban kapott (a modellből számított) formánsfrekvenciák értékének pontosságával kapcsolatban aggályaik merültek fel – úgy vélték ugyanis, hogy az általuk mért formánsfrekvenciák a feltételezett valós frekvenciáknál (a JOHANNSSON és munkatársai (1982) végezte vizsgálathoz hasonlóan) alacsonyabbak voltak.

Az énekelt magánhangzók formánsait kizárólag akusztikai adatok alapján elemző vizsgálatok egyik és talán legfontosabb eltérése a hagyományos akusztikai elemzésekhez képest egy, a beszélőtől független, a természetes fonációt helyettesítő külső gerjesztés, külső forrásjel alkalmazása. Az itt bemutatásra kerülő vizsgálatok nagy részében ugyanis a kutatók egy nagy felhangsűrűségű, alacsony alapfrekvenciájú külső gerjesztésű jelet juttattak az artikulációs csatornába (technikától függően a nyaknál, a gége fölött vagy a szájon keresztül), majd ezt a toldalékcsőből kijutva, a szájnyílásnál (mikrofon segítségével) rögzítették. A külső gerjesztésű, de a toldalékcső által megszűrt hang elemzésével ugyanis az akusztikai kimenetben is „láthatóvá váltak” a magas alapfrekvencián ejtett hangok formánsfrekvenciái, melyek a természetes hangadásban a magas alapfrekvencia és a felhangok közti nagy távolság miatt nem állapíthatók meg a hagyományos, a kiejtett beszédhangnak a Fourier-transzformáción alapuló elemzésével. A külső forrásjel tehát lehetővé teszi, hogy a toldalékcső átviteli karakterisztikáját a fonációtól független, alacsony alapfrekvenciájú, nagy felhangsűrűségű hang segítségével „újra mintavételezzük”, és ennek a jelnek az elemzésével kikerülhetővé válik a magas alapfrekvenciájú hangforrás elemzésének problémája.

SUNDBERG (1975) már említett vizsgálatában nem csak artikulációs, hanem akusztikai méréseket is végzett: a svéd /u o a e i y/ hangzókat elemezte, 262 Hz és 698 Hz alapfrekvenciák (a zenei c' és g'' hang) között egy szoprán ejtésében, úgy, hogy az egyes hangok ejtésekor az énekes zöngéjét egy elektromos gége gerjesztette külső forrásjellel helyettesítette. A felvételek a következőképpen zajlottak. Az énekes először saját hangszalagrezgésének segítségével produkálta az énekelt magánhangzókat, majd bekapcsolt egy, a gégeje felett (a garat

magasságában) a nyakához illesztett elektromos géget, mely így mély alaphangú, nagy felhangsűrűségű berregő hangot juttatott a toldalékcső alsó részébe. Ezután egy adott pillantaban az énekes abbahagyta a zöngékepzést (ekkor már csak az elektromos gége keltette berregő hang jelentette a száját elhagyó kimeneti jel forrását), és bezárta a hangrést (a levegő visszatartásával). Mivel a fonáció abbahagyásakor az énekes igyekezett megmerevíteni az artikulációját (és még pár másodpercig „némán énekelve” tartani azt), az ekkor a száját elhagyó nagy felhangsűrűségű, az elektromos gége keltette akusztikai jelben az artikulációs konfiguráció aktuális sajátfrekvenciái emelkedtek ki, és váltak mérhetővé (a szerző tehát az elektromos gége alapvető működési elvét használta ki). (Erről a módszerről bővebben lásd még a 6.4.1. *alfejezetet*, ahol az elektromos gége használatát, valamint az elektromos gége használatával kapott akusztikus jelet a jelen kötetben is bemutatott vizsgálatok tükrében illusztráljuk).

A vizsgálatában kapott rezonanciaadatok segítségével SUNDBERG (1975) bizonyítékkal szolgált az $F_1 : f_0$ hangolásnak a hangzók nyíltsági fokától, illetve első formánsának értékétől függő tendenciájára a kísérletben vizsgált /u o a e i y/ magánhangzók esetében. Ennek megfelelően SUNDBERG azt is megállapította, hogy az /a/ formánsai a többi magánhangzónál jóval nagyobb függetlenséget mutattak az alapfrekvenciától, SUNDBERG feltételezése szerint éppen azért, mert az /a/ nyílt ejtésű, és magasabb első formánssal rendelkezik.

HERTEGÁRD és GAUFFIN (1993) akusztikai vizsgálata csak az /a/ hangra koncentrált 250, 390 és 750 Hz alapfrekvenciákon. Vizsgálatuk során a kutatók a résztvevő szoprán énekes hangját először egy Rothenberg-maszk, majd pedig egy mikrofon segítségével rögzítették. A Rothenberg-maszk a hagyományos mikrofonhoz képest nem a hangnyomás-idő függvényt, hanem a beszéd közben a száját elhagyó levegő mennyiségét (illetve a levegő térfogatsebességét) regisztrálja egy úgynevezett áramlási glottogramban. Az áramlási glottogram hullámformája tehát a levegő áramlását milliméter/másodperc (mm/s) mennyiségként rögzíti az idő függvényében. Az elemzések során HERTEGÁRD és GAUFFIN a két rögzített hullámforma, azaz a hangnyomás-idő függvény és az áramlási glottogram inverz szűrését végezte el. Mivel az inverz szűrés során az elemző arra törekszik, hogy „kioltva” az áramlási függvényben az üregi rezonátortevékenység okozta energiamaximumokat „rekonstruálja” a transzglottális levegő-áramra még a glottisznál jellemző áramlási (térfogat-sebességi) függvényt (bővebben lásd még a jelen értekezés 6.2. *alfejezetét*), a szűrés során egyúttal a toldalékcső szűrőtevékenységét jellemző rezonanciahelyeket is megállapítja. Ezt használta ki HERTEGÁRD és GAUFFIN (1993) is: az inverz szűréssel a szerzők nem elsősorban a hangrésnél jellemző áramlási glottogramot, hanem a glottogramot az éneklés során befolyásoló (megszűrő) toldalékcsővi sajátrezonanciákat, azaz a formánsfrekvenciákat becsülték meg.

Eredményeik szerint a szoprán énekes az alaphang emelésével növelte az F_1 értékét úgy, hogy az megközelítse az f_0 frekvenciaértékét akkor, ha az f_0 az F_1 beszédben mérhető értékénél magasabbra emelkedett (az énekes tehát hangolta az F_1 formánst). Ugyanakkor a szerzők azt is kimutatták, hogy a mikrofonnal rögzített hanganyagon a Rothenberg-maszkos felvételekből detektálhatóknál magasabb F_1 értékek voltak mérhetőek. A két jel szűrésének eredményeként előállt formánsfrekvenciabeli különbségeket azzal magyarázták, hogy a maszki (zárt

formájából adódóan) korlátozta az énekest állkapcsa szabad mozgásában (azaz az állkapocs szabad nyitásában), mely feltételezésüket, leírásuk szerint, az énekes beszámolója is megerősítette.

JOLIVEAU és munkatársai (2004) is külső gerjesztésű forrásjelet alkalmaztak akusztikai mérésükben, ám ők ezt a forrásjelet az énekes szájnyílásánál juttatták a toldalékcsőbe, úgy, hogy besugárzás közben az énekes folyamatosan énekelt. A felvételek során a toldalékcsőből visszaverődött jelet egy mikrofon segítségével rögzítették, majd „megtisztították” az énekes fonációja gerjesztette hangtól, és végül akusztikailag elemezték – azaz meghatározták a jelbesugárzás alatt jellemző artikulációs konfiguráció formánsfrekvenciáit. A szerzők ezzel a módszerrel vizsgálták az ausztrál angol /a ɔ u ɜ/ magánhangzókat (melyeket a *hard, hoard, who'd, heard* hangsorokkal elicitáltak). A kilenc szoprán résztvevő halkán (piano módban) és vibrató nélkül (senza vibrato) énekelt a vizsgált hangzókat kitarva (körülbelül 4 s hosszúságban), az egyes énekesek számára kényelmes alapfrekvenciákon (maximálisan körülbelül 1046 Hz-en). A szerzők a vizsgálatban kimutatták az F_1 -nek¹⁴ az f_0 -ra hangolását olyan esetekben, amikor az f_0 megközelítette az F_1 beszédben mért értékét. Azt találták azonban, hogy bár a hangolási tendencia még megközelítőleg 1000 Hz alapfrekvenciáig jellemző volt az ajakréses /a/ és /ɜ/ magánhangzókra, az ajakkerekítéses /ɔ/ és /u/ esetében 1000 Hz alapfrekvencián már nem jelent meg (itt az F_1 az f_0 -nál alacsonyabb értéken volt mérhető). JOLIVEAU és munkatársai (2004) az előbbieken felül ráadásul azt is feljegyezték, hogy az F_1 hangolásának megjelenésével az F_2 értéke is változott, de a szerzők értelmezése szerint ezt pusztán az F_1 hangolásából következett (sőt ez annak egyenesen elkerülhetetlen velejárója volt). A szerzők az F_2 -re nézvést tehát nem feltételeztek szándékos hangolási stratégiát, és arra következtettek, hogy az F_1 emelésének eléréséhez szükséges artikulációs változtatások szükségszerűen az F_2 értékére is hatással vannak.

GARNIER és munkatársai (2010) ugyanezzel a szájnyílásnál a toldalékcsőbe juttatott külső gerjesztést alkalmazó módszertani megoldással elemezték az /a/ magánhangzót 440 Hz alapfrekvencia fölött 12 szoprán énekes (négy profi, négy haladó és négy kezdő) ejtésében. Eredményeik szerint az énekesek az f_0 értékéhez hangolták az /a/ első formánsát 1000-1200 Hz (a c''' és d''' zenei hangok) alatt a képzettségi fokuktól nagyjából függetlenül. 1200-1400 Hz fölött (2300 Hz-ig) azonban az $F_1 : f_0$ hangoláson felül egy további hangolási tendenciát is találtak. Ezen a kifejezetten magas alapfrekvencia-tartományon ugyanis öt énekes nem az F_1 , hanem az F_2 értékét emelte az f_0 értékére (vagy annak közelébe) – míg volt, aki továbbra is az $F_1 : f_0$ hangolást alkalmazta. A szerzők az 1200 Hz fölötti extrém alapfrekvenciákra talált eredményeikkel, illetve az $F_2 : f_0$ hangolási stratégia létezésével tartják magyarázhatónak azt,

¹⁴ Megemlítendő, hogy az idézett tanulmány élesen elkülöníti az üregi rezonanciákat a fonáció mintavételezte kimeneti jel intenzitásmaximumaitól. Ennek értelmében az előbbiekre az R_1 , R_2 stb. notációt, míg az utóbbiakra az F_1 , F_2 stb. notációt alkalmazza. Mivel a jelen értekezésben nem követjük ezt a jelölési konvenciót, és formánsnak az üregi rezonanciákat tekintjük (akkor is, ha azok nem esnek egybe a kimeneti jel intenzitásmaximumaival), ezért mind JOLIVEAU és munkatársainak (2004), mind pedig GARNIER és munkatársainak (2010) a rezonanciákra vonatkozó eredményeit F_1 , F_2 stb. megjelöléssel hivatkozunk.

hogy a szopránok képesek az ezen a frekvencián már igen kis intenzitású zöngével is a kívánalmaknak megfelelően nagy erejű hangadásra.

Összegzésképpen tehát az ismertetett vizsgálatok a svéd és az ausztrál angol nyelv bizonyos magánhangzóit érintették, illetve több esetben csak az /a/ hangzót a nyelv megjelölése nélkül. A vizsgálatok fő kérdése a formánshangolási stratégiák megjelenése volt, különös tekintettel az F_1 -re. A vizsgálatokból származó (direkt és indirekt akusztikai, illetve artikulációs eredmények) gyakorlatilag egybehangzóan kimutatták az F_1 -nek az alapfrekvenciára hangolását akkor, amikor az f_0 emelésével annak értéke megközelítette az F_1 -nek a beszédben az adott magánhangzóra jellemző frekvenciaértékét. Két tanulmányban kivételként említettek a szerzők az /i u/ magánhangzókat: SUNDBERG és SKOOG (1995) ezek esetében ugyanis nem talált az F_1 emelésére utaló állkapocsnyitási módosításokat, míg BRESCH és NARAYANAN (2010) éppen e hangzónál mutatta ki legegyszerűbben az $F_1 : f_0$ hangolást. Az eredmények alapján leszűrhető további fontos és hangsúlyozandó tanulság, hogy az akusztikai adatok azt is mutatják, hogy a szoprán énekes által ejtett magánhangzók minősége mind az ejteni szándékozott „eredeti” magánhangzó-minőségnek, mind pedig az alapfrekvenciának a függvényében, az alapfrekvencia emelésével párhuzamosan fokozatosan változik. Mindebből következően tehát felvethető (és feltehetőleg egyáltalán nem triviális) kérdés, hogy ezek a tipikus megvalósulásukhoz képest többé-kevésbé eltérő, magas alapfrekvencián énekelt magánhangzók miként sorolhatók be az egyes az adott nyelvre jellemző fonémakategóriák valamelyikébe, azaz milyen beszédhangokként és milyen hatékonysággal azonosíthatók a hallgatók számára.

A fent említett vizsgálatok módszertani bemutatásában a szerzők az alkalmazott módszerekben rejlő hibalehetőségeket is kiemelték, valamint rávilágítottak, hogy az egyes módszerek a megfigyelni kívánt artikulációs-akusztikai stratégiák módosítására készítették az énekeseket. A módszertani problémák közt legfontosabbként az artikulációs adatokkal dolgozó vizsgálatok gyakorlatilag mindegyike a keresztmetszetfüggvény-számításban, valamint a fizikai vagy matematikai modellben végzett rezonanciamérésekben rejlő potenciális mérési pontatlanságokat említette (BRESCH és NARAYANAN (2010) éppen ezért például egyenesen kerülte a konkrét formánsértékek bemutatását és elemzését, csak statisztikailag vetette össze őket az f_0 -értékekkel összehasonlítva). A Rothenberg-maszkot használó akusztikai elemzésen alapuló vizsgálatban a kutatók a maszk akadályozó szerepét emelték ki, ugyanis a maszk alsó pereme, mint írják, korlátozta az énekest az állkapocsnyitás szabad növelésében (HERTEGÁRD–GAUFFIN 1993). JOLIVEAU és munkatársai (2004), valamint GARNIER és munkatársai (2010) leírják, hogy vizsgálatukban az énekest halk és lehetőség szerint teljesen vibrátómentes énekmodra kérték (a mérés technikai megoldás érzékenysége miatt), ez azonban mind a magasabb alapfrekvenciák produkciós szükségszerűségeit, mind pedig az énekprodukción meg szokott módját, természetességét tekintve problematikus lehet. Egyrészt azért, mert az f_0 emeléséhez köztudott módon természetszerűleg nagyobb szubglottális nyomás és nagyobb hangerősség szükséges, másrészt pedig azért, mert a szólóénekes általánosságban vibrátót alkalmazva gyakorol és ad elő. Éppen ezért feltételezhetjük, hogy JOLIVEAU és munkatársai

nak (2004), valamint GARNIER és munkatársainak (2010) vizsgálata nem az énekes által leginkább begyakorlott technikával (és feltehetően nem az operai kívánalmaknak leginkább megfelelő módon) megvalósított énekhang elemzése alapján készült, továbbá azt is valószínűsíthetjük, hogy az énekhang természetellenes halkítása és a vibrató nélküli megvalósítása némileg befolyásolta az énekes által alkalmazott formánsstratégiákat is. SUNDBERG és munkatársainak (2013) kritikai észrevétele szerint ráadásul a szájnál, fonáció közben bejuttatott külső gerjesztésen alapuló mérés azért is problémás, mert így a fonációs ciklus azon szakaszában, mely során a glottisz nyitva van, nem kívánatos módon a szubglottális üregrendszer rezonátortevékenysége is megjelenik a kimeneti jelben.

A fent említett vizsgálatok javában azt is láttuk, hogy azok vagy egy hangzóra (tipikusan az /a/-ra) koncentráltak, vagy az artikulációs-akusztikai magánhangzótér szélső értékeit képező /a i u/ hangokra – a legtöbb esetben tehát nem vizsgálták a nyelvspecifikus magánhangzórendszer hatását az egyes magánhangzók megvalósulására. Esetenként, ha a vizsgálatban ezeknél több hangzómínőség szerepelt, a hangzók kiválasztásának oka még mindig nem egy-egy nyelv magánhangzórendszerének a teljes feltérképezése volt, inkább csak az egyes eltérő nyelvállásfokok mint képzési jegyek reprezentációja. Az eddigi akusztikai vizsgálatok többsége tehát nem szentelt különösebb figyelmet a magánhangzók artikulációs-akusztikai változásaiban fellelhető nyelvspecifikus tendenciáknak, illetve a teljes magánhangzótér alakulására jellemző sajátosságoknak.

Az említett akusztikai vizsgálatokban egy másik, méltatlanul mellőzött aspektus (és amint azt a következő alfejezetben látni fogjuk, ez a szempont a magánhangzók észlelésével foglalkozó percepció vizsgálatokból is hiányzik) az artikulációs-akusztikai eredmények és a percepció tendenciák közti összefüggések részletekbe menő elemzése és az akusztikai-artikulációs változásokat követő percepció tendenciák nyelvspecifikus leírása. E mellőzés oka egyfelől persze nagyon is jól érthető, hiszen az alacsony alapfrekvenciájú beszédprodukciónak a percepciójára kidolgozott modellek sem az $F_1 : f_0$ formánshangolás, sem pedig a magas alapfrekvencia okozta alacsony kimeneti felbontás okán nem adnak egyértelmű vagy könnyen igazolható predikciókat. Azonban azt is látnunk kell, hogy egyes kutatók hajlamosak gyakorlatilag teljesen elvitatni a kérdés komplexitását (és sokszor a relevanciáját is), és triviális kérdésnek tekintve azt átsiklanak fölötte egy-két, az akusztikumból származó egyszerű következtetés után. JOLIVEAU és munkatársai (2004) például azt a predikciót fogalmazzák meg, hogy mivel a spektrális alul-mintavételezettség miatt már amúgy is egyértelműen csökken a magánhangzók „érthetősége” magas alapfrekvencián, nem meglepő, hogy az énekesek a hangzómínőségek megtartása helyett a hangerősség és hangszínezet produkciójára koncentrálnak a formánsok értékének változtatását (azaz az $F_1 : f_0$ hangolást) választják stratégiaként. Értelmezésük szerint tehát a magánhangzók érthetőségének csökkenése már pusztán a spektrális alul-mintavételezettség fokozódásából következően is szükségszerű (és egyértelmű), amin az F_1 hangolása már „nem ront sokat”, így – implicit következtetésük szerint – a két faktor együttes hatása már relatíve alacsony alapfrekvenciákon is gyakorlatilag felismerhetetlenné teszi a magánhangzókat. Ám a valós helyzet – amint azt az észlelést vizsgáló

tanulmányok, valamint a saját, itt bemutatásra kerülő kísérleteink tárgyalásakor látni fogjuk – jóval komplexebb ennél.

Hangsúlyozni kívánjuk, hogy a rezonanciaadatok és a percepciók adatok összefüggésbe hozásával kapcsolatos aggályokat és elvi-gyakorlati nehézségeket ismerjük, és érthetőnek tartjuk, hiszen a magánhangzó-észlelés hagyományos, formánsokon és azok viszonyán alapuló értelmezése a magas alapfrekvenciák esetében valóban nehezen értelmezhető, vagy egyenesen akár értelmezhetetlen is lehet. Ugyanakkor úgy véljük, hogy nem lenne haszon nélkül való, ha szisztematikus áttekintést nyerhetnénk az akusztikai és percepciók adatok közti lehetséges összefüggésekről egy nyelv (lehetőség szerint) teljes magánhangzókészletének vizsgálata segítségével – és persze minden, a magas alapfrekvenciából és a formánshangolás jelenségéből következő megszorítást szem előtt tartva.

A következőkben a magas alapfrekvencián énekelt magánhangzók észlelésének kérdésében eddig folytatott vizsgálatok eredményeit ismeretjük.

1.2.3. Az énekelt magánhangzók észlelése

Az éneklés produkciós sajátosságait vizsgáló kutatások alapján feltételezhető, hogy az énekelve ejtett magánhangzók eltérése a mindennapi beszédben megvalósuló magánhangzókhoz képest jóval nagyobb lehet, mint a magánhangzókra a hangkörnyezettől és a beszélőtől függő variabilitás következtében a beszédben jellemző változatosság. Az éneklés ilyen szempontból tehát a beszéd határhelyzetének, extremitásának tekinthető – úgy értelmezhető, mint egy olyan beszédprodukciós forma, mely egyfajta „természetes” módon feszegeti a beszédhangok megvalósításának határait. A különleges produkciós mód révén egyúttal a beszédpercepció is új oldaláról válik vizsgálhatóvá, ha ezen extrémnek tűnő beszédhang-megvalósítások észleleti vetületét elemezzük.

Noha úgy tekinthetjük, hogy az éneklés például a laboratóriumi körülmények között felvett és akár gépileg manipulált beszéddel szemben a beszédprodukció valamely természetes extremitását képezi, nélkülözi a laboratóriumi beszéd vizsgálatának egyes nagyon hasznos sajátosságait is. Az énekelt magánhangzók vizsgálatában ugyanis bizonyos jelenségek – legfőképpen az alapfrekvencia és a formánsfrekvenciák alapfrekvencia-függő változásai – nem választhatók szét egymástól, hiszen az éneklésben nincsen egyik jelenség a másik nélkül, nem volna tehát az éneklésre nézvést értelmezhető egy, a jelenségek disszociációján alapuló kísérlet. De ragadjuk meg a kérdést mégis inkább a másik oldaláról, és máris optimistább képet kapunk az énekelt beszédhangok vizsgálatának hasznosságáról. Éppen abban rejlik ugyanis az énekelt magánhangzók érdekessége, hogy azok egyes olyan jelenségeknek a természetes ötvözetét képezik a mindennapokban tapasztalható beszédhang-megvalósításban, melyekről a percepciók szakirodalom alapján tudjuk, hogy befolyásolják, sőt alapvetően meghatározzák az egyes magánhangzó-minőségeket, azok észlelését. A jelenségek az éneklésben tapasztalható komplex érvényesülésének észleleti hatása azonban egyáltalán nem magától

értetődő, azaz nem válaszolható meg pusztán a már ismert beszédpercepciók modellek valamelyikének egyszerű adaptációjával, az énekelt magánhangzók észlelését illető kérdések megválaszolásához kísérleti bizonyítékok szükségesek.

A következőkben előbb ismertetjük, milyen predikciókat fogalmaznak meg az egyes, az akusztikummal foglalkozó tanulmányok az énekelt magánhangzók azonosítására és azonosíthatóságára, utóbb pedig bemutatjuk a kérdés eddigi kísérletes eredményeit.

1.2.3.1. AZ ÉNEKELT MAGÁNHANGZÓK ÉSZLELÉSÉVEL KAPCSOLATOS GYAKORI FELTÉTELEZÉSEK

Az énekelt magánhangzók artikulációs és akusztikai szerkezetének ismeretében számos predikció megfogalmazható az énekelt magánhangzók észlelésére. Látni fogjuk azonban, hogy ezen feltételezéseknek csak egy elenyésző részével kapcsolatban áll rendelkezésünkre megbízható kísérletes bizonyíték.

Az első az énekelt magánhangzók észlelését érintő, de kísérleti bizonyítékok híján megfogalmazott feltételezés (melyet például JOLIVEAU és munkatársai 2004 tanulmányában is olvashatunk) a következő. A magánhangzók magas alapfrekvenciákon énekelve az alapfrekvencia emelésével minőségükben egymáshoz egyre hasonlóbba lesznek, míg végül kellően magas alapfrekvencián teljesen uniformizálódnak. Ez az uniformizálódás ráadásul – szól a másik fontos jóslat, melyet ugyanitt olvashatunk – úgy megy végbe, hogy az f_0 emelését kísérő egyre növekvő állkapocsnyitásszöggel és egyre magasabb F_1 -gyel ejtett magánhangzók egyre nyíltabb magánhangzókként lesznek felismerhetők. Hangsúlyozandó, hogy az idézett szerzők ez utóbbi felvetést annak ellenére fogalmazzák meg, hogy a magánhangzók nyíltságának észleletéről már több vizsgálat is bizonyította, hogy elsősorban nem az F_1 abszolút értékének, hanem annak az f_0 -hoz viszonyított tonotopikus távolságának függvénye (vö. 1.1.2.2. *alfejezet*). Egy harmadik gyakori feltételezés szerint a mássalhangzó-környezet a koartikulációs hatásokon keresztül minden bizonnyal „javít” az ejtési szándék szerinti magánhangzó-azonosításon még a magas alapfrekvencián énekelt magánhangzók esetében is (vö. pl. SCOTTO DI CARLO–GERMAIN 1985; GOTTFRIED–CHEW 1986; HOLLIEN et al. 2000). Az ezt az állítást megfogalmazó szerzők azonban az esetek többségében nem foglalnak egyértelműen állást abban a kérdésben, hogy a mássalhangzók hatását a percepciók vagy az artikulációs-akusztikai vetületben feltételezik-e, miközben pedig az egyes magánhangzók az f_0 függvényében változó minőségét maguk is demonstrálják (erről az egyes percepciók tanulmányok tárgyalásánál több szót ejtünk).

Amint azt a következő alfejezetben látni fogjuk, mindhárom predikció esetében elmondható, hogy azok inkább csak spekulatív, de nem kísérletekkel ellenőrzött feltevéseken alapulnak. A specifikusan e kérdések vizsgálatát célzó, nagyon szisztematikus és jól kontrollált kísérletek ugyanis a mai napig hiányoznak a szakirodalomból.

1.2.3.2. AZ ÉNEKELT MAGÁNHANGZÓK ÉSZLELÉSÉNEK VIZSGÁLATA

SUNDBERG 1975-ben született tanulmánya óta kísérletes bizonyítékokkal rendelkezünk arról, hogy az énekelt magánhangzók akusztikuma az alapfrekvencia és a magánhangzó-minőség függvényében többé vagy kevésbé eltér a beszédben ejtett magánhangzók akusztikai szerkezetétől. Amint azt az 1.2.2. *alfejezetben* bemutattuk, SUNDBERG (1975) és az őt követő szerzők artikulációs és akusztikai elemzéseinek értelmében a magas alapfrekvencián énekelt magánhangzók és beszédbeli megvalósulásuk közti legfőbb különbség az első formánsnak az alapfrekvencia értékére (vagy valamivel a fölé) hangolása azokban az esetekben, amikor a megemelt alapfrekvencia magasabb lenne a magánhangzóra a beszédben jellemző első formáns frekvenciaértékénél. Az említett alfejezetben ezen túlmenően azt is kiemeltük, hogy a beszéd és a magas alapfrekvenciájú éneklés között további jelentős eltérések származnak a magas alapfrekvencia okozta spektrális alul-mintavételezettség jelenségéből is, azaz abból a pusztán alapfrekvencia-függő jelenségből, melynek eredményeképpen a magas alapfrekvenciákon ejtett magánhangzók felbontása alacsonyabb az akusztikai kimenetben a mélyebb alapfrekvenciákon ejtett magánhangzókéhoz képest.

Mіндеzen felfedezések eredményeképpen az idézett SUNDBERG-tanulmány (1975) megjelelése óta az is egyre terjedő nézetnek számít a szakirodalomban, hogy a magas alapfrekvenciájú (szoprán) éneklésben ejtett magánhangzók „érthetősége” (az angol *intelligibility* szóból fordítva) vagy észlelhetősége, azaz e magánhangzóknak az ejtési szándék szerinti azonosítása is alacsony. Valamivel SUNDBERG (1975) vizsgálata után ezt a sok esetben kísérletes igazolás nélkül elfogadott feltételezést több vizsgálat is górcső alá vette. E vizsgálatok központi kérdéseit a magas alapfrekvencián énekelt magánhangzókra az alapfrekvencia függvényében jellemző azonosítási tendenciák, valamint az azonosítást segítő vagy éppen akadályozó körülmények (pl. a fonetikai kontextus vagy annak hiánya) képezték. Az elkövetkezőkben ezekről a vizsgálatokról lesz szó.

Mielőtt azonban rátérnénk az énekelt magánhangzók kísérletes vizsgálatának eddigi eredményeire, fontosnak tartjuk e helyütt egy alapvető fogalmi kérdés tisztázását, mely feltehetőleg a korábbi szakirodalomban honoló ellentmondások egy jelentős részéért felelőssé tehető, és – részben éppen ezért – a jelen kötetben központi szereppel bír. A nemzetközi (angol nyelvű), az énekelt magánhangzók azonosításával foglalkozó szakirodalomban az eredmények értékelésénél általánosan azt találjuk, hogy a válaszként megjelent magánhangzókat az ejteni szándékozott magánhangzóhoz viszonyítva „helyes” (*correct*) vagy „hibás/téves” (*incorrect/mistake/error*) válaszokként minősítik a szerzők. Teszik ezt annak ellenére, hogy a legtöbb esetben az adott vizsgálatban nem térnek ki az ejtett hangzók akusztikai szerkezetének elemzésére, pusztán az ejteni szándékozott magánhangzókat listázzák, és ezekhez viszonyítják a válaszokat. A „helyes válasz” és a „téves válasz” megfogalmazások tehát erősen félrevezetők, hiszen azt sugallják, hogy a válaszok kiértékelésekor pontosan tisztában vagyunk az adott beszédhangoknak az akusztikai szerkezet által definált (produkciós) minőségével, azaz a „helyes válasszal”. Mindez azonban mind SUNDBERG (1975) eredményeinek,

mind pedig minden későbbi, az énekelt magánhangzók akusztikai szerkezetét vizsgáló eredménynek ellentmond (vö. előző alfejezet), hiszen ezek a vizsgálatok éppen a magánhangzók-nak a hangzómínőség és az alapfrekvencia függvényében az alapfrekvencia növelésével egyre inkább megváltozó akusztikai szerkezetéről tanúskodtak.

Az előző alfejezetben ismertetett produkciós vizsgálatok egyik legfontosabb – de a szakirodalomban talán nem kellően hangsúlyozott – tanulságának éppen azt tartjuk, hogy azok eredményei szerint a magánhangzók realizációja az éneklésben az adott magánhangzó minőségének és az alapfrekvenciának a függvényében az alapfrekvencia emelésével megközelítőleg fokozatosan változik, egyfelől a hangzók fokozódó spektrális alul-mintavételezettsége, másfelől pedig a hangzók formánsszerkezetének változása miatt. Ennek értelmében tehát a magas alapfrekvenciákon énekelt magánhangzók minősége, a magánhangzó-minőségek fonemikus besorolása nem triviális, sőt vélekedésünk szerint a fonemikus besorolást még akusztikai adatok birtokában sem tekinthetjük egyértelműnek vagy adottnak. A fent említett változások ugyanis – az eddigi eredmények alapján – alapvetően nem kategóriaváltásszerűek, azaz nem egy, az adott nyelven létező beszédhang egy másik, az adott nyelvben létező beszédhangként való megvalósításaként írhatók le, hanem graduálisak. Ez viszont azt jelenti, hogy a megváltozott akusztikai szerkezetű hangok percepciók vetülete éppen a beszédpercepció természetéből fakadóan nem lehet triviális, és e módosult akusztikai szerkezetű hangok besorolása csak kísérletes úton állapítható meg.

A fentiek értelmében úgy gondoljuk, hogy az énekelt magánhangzók percepciók (és ezzel kiegészülő akusztikai) vizsgálatának egyik legfontosabb célja az alapfrekvencia változásának következtében módosult akusztikai szerkezetű hangok percepciók vetületének elemzése. A tesztekben kapott válaszok „helyes” és „téves” minősítésével tehát éppen a vizsgálatban megcélzott kérdést „válaszolnánk meg” előre (ráadásul sok esetben valószínűleg tévesen). Ennek megfelelően az értekezésben lehetőség szerint kerüljük a magánhangzók azonosításának „helyes” vagy „téves” minősítéseit, helyette a magánhangzók az **ejtési szándék szerinti** vagy attól eltérő (fonemikus) **besorolásáról**, a magánhangzók azonosításáról beszélünk, így kerülve el azt a véleményünk szerint értelmezhetetlen állásfoglalást, mely szerint a kísérletekben részt vevő hallgatóknak egy adott magánhangzót kellett volna felismerniük a kategória határokat feszegető, nem egyértelmű akusztikai szerkezetű hangokban. Vegyük észre azonban, hogy még valamelyest az általunk használt megjelölés is félrevezető lehet, ha arra gondolunk, hogy az ejtési szándék szigorúan véve nem csak az adott feladatban az énekes előtt szereplő fonemalista vagy szólista megvalósításának szándékát, de már az adott esetben, az adott alapfrekvencián esetleg szükséges formánshangolás (vagy más, az éneklés során elsajátított formánssztratégiák) alkalmazásának intencióját is magában foglalja. Zavarónak tartanánk azonban egy még ennél is pontosabb, így szükségszerűen komplexebb, terjedősebb megfogalmazás használatát az értekezésben. Éppen ezért megmaradunk az „ejtési szándék szerinti azonosítás” formulánál, azzal a kitételrel, hogy az „ejtési szándékot” kizárólag az énekes adatközlőknek a felvételek alkalmával mutatott (stimulus-, vagy fonemalista által megjelölt) magánhangzó-minőségek megvalósításának szándékára értjük, nem pedig az egyes

magánhangzóknek az énektechnika elsajátítása során tanult, az alulfrekvenciától függően módosított megvalósításának stratégiájára.

A következőkben az énekelt magánhangzók azonosítására vonatkozó szakirodalom legtöbbet hivatkozott, azaz legnagyobb befolyásúnak tűnő tanulmányait ismertetjük.

SMITH és SCOTT (1980) azt a kérdést igyekezett körüljárni, hogy lehetséges-e – és ha igen, miként – a magas alulfrekvencián énekelt magánhangzókat érintő akusztikai változásoknak (de főként az $F_1 : f_0$ formánshangolásnak) a kiküszöbölése az észleletben. Kísérletükben a szerzők abból az alapfeltevésekből indultak ki, hogy a magas alulfrekvencián énekelt magánhangzók akusztikai szerkezete megváltozik, és ebből egyenesen következik az is, hogy az ejteni szándékozott magánhangzó az ejtési szándék szerinti azonosításának aránya egyre csökken az alulfrekvencia emelésével. Az $F_1 : f_0$ formánshangolásról bevezetőjük értelmében a szerzők azt feltételezték, hogy az elsősorban az alacsony első formánssal rendelkező, azaz a zárt, első nyíltsági fokú, felső nyelvvállású magánhangzók éneklésekor fellépő jelenség, így elsősorban e hangzók esetében jelenthet az észleleti magánhangzó-minőséget érintő problémát. SMITH és SCOTT (1980) STRANGE és munkatársainak (1976) kísérleti eredményeire építve azt a feltételezést igyekezett megvizsgálni, mely szerint a magánhangzók az ejtési szándék szerinti azonosítása hatékonyabb, ha a magánhangzót mássalhangzó-környezetben ejtve halljuk. (Emlékeztetőül, STRANGE és munkatársai (1976) vizsgálataikban azt igazolták, hogy a magánhangzót körülvevő mássalhangzók a magánhangzó és a mássalhangzók közötti hangátmenetekben rejlő akusztikai információon keresztül segítik a magánhangzó felismerését akár értelmetlen szavakban is, sőt akár még olyan helyzetekben is, amikor a teljes magánhangzós rész törlődött a CVC szekvenciából, vö. a jelen értekezés 1.1.2.2. *alfejezetét*).

Mivel egy, az énektanítás elméletének történetében már nagyon régóta létező feltételezés (vö. GARCIA 1911) – illetve a SMITH és SCOTT (1980) vizsgálatát megelőző egyes kísérletes eredmények – szerint a jó szoprán énektechnika a gége lesüllyesztését alkalmazza (vö. a jelen értekezés 1.2.1. *alfejezetét*), a szerzők a mássalhangzó-környezet hatásán kívül a gégehelyzet hatását is tesztelni kívánták percepció vizsgálatokban.

A kísérlet anyagát az angol /i ɪ ε æ/ magánhangzók képezték a *bVd* hangkörnyezetben (azaz a *bead* 'gyöngy', *bid* 'ajánlat, tét', *bed* 'ágy', *bad* 'rossz' minimális négyesben, „CVC” kondíció) és izolációban („V” kondíció), az *f*' (349 Hz), *a*' (440 Hz), *c*#" (554 Hz), *f*#" (698 Hz), *a*#" (880 Hz) és *c*#" (1108 Hz) alulfrekvenciákon, egy szoprán énekes énekprodukción. A „V” kondícióba tartozó magánhangzókat a szerzők izolált ejtésben vizsgálták, azaz fokozatos zöngéindítást, hangindítást vagy zöngékezdetet (onset) és fokozatos zöngécsengést (offset) tartalmazó hangokban. A két, a gégehelyzettel összefüggő kondíciót a szerzők „operai” és „megemelt gége” énekprodukciónak nevezték, és bár ezek elicitálásának módjáról és kontrollálásáról nem adtak pontosabb leírást, a módokat valamelyest definiálták tanulmányukban. Meghatározásuk szerint az „operai” produkciós mód (a „megemelt gégevel” szemben) a „rendes” operaéneklésben használt technika, melynek alkalmazását a szerzők – feltehetőleg – auditíve (saját hallásuk segítségével) ellenőrizték. A percepció tesztet a szerzők 10 résztvevővel (egyetemi hallgatóval) végezték el, az egyes magánhangzókat

randomizált, de az egyes alapfrekvenciákat frekvenciájuk szerint egymást követő, növekvő sorrendben játszották le a kísérlet során.

SMITH és SCOTT eredményei alapján a legmagasabb (1108 Hz) alapfrekvencián az ejtési szándéknak megfelelő azonosítás szerint az egyes kondíciók közt tapasztalt sorrend a következő volt: „megemelt gége” + „CVC” (76%) > „operai” + „CVC” (70%) > „megemelt gége” + „V” (44%) > „operai” + „V” (12%). Alacsonyabb frekvenciákon a kondíciók itt bemutatott sorrendjében a szerzők variabilitást találtak, mert például a „megemelt gége” + „V” kombináció sok esetben jóval, akár 32%-kal is magasabb százalékos arányokat mutatott, mint az „operai” + „CVC” kombináció. Az „operai” + „V” kombináció minden alapfrekvencián elmaradt a többi kombinációhoz képest, míg a „megemelt gége” + „CVC” kombináció gyakorlatilag minden alapfrekvencián a legmagasabb azonosítási százalékokat mutatta (azaz itt egyezett a legnagyobb számban az ejtési szándék szerinti magánhangzó-minőség az azonosított magánhangzóval). Az „operai” mód a „V” és „CVC” kondícióban kapott eredményei a magas (f'' vagy 698 Hz-es és a feletti) alapfrekvenciákon szisztematikusak voltak, és a mássalhangzó-környezetben ejtett hangok előnyét mutatták az izolált ejtéssel szemben (rendre: 16% és 54%, 4% és 68%, valamint 12% és 70%). SMITH és SCOTT a f'' (698 Hz) feletti, tehát az általuk magasnak tekintett alapfrekvenciák kiátlagolásával a következő sorrendet állították fel a kondíciópárok között: „megemelt gége” + „CVC” (83%) > „operai” + „CVC” (64%) > „megemelt gége” + „V” (62%) > „operai” + „V” (10%).

Eredményeiket a szerzők a következőképpen foglalták össze. 1. A magas alapfrekvencián operai módban énekelt, izolált ejtésű magánhangzók azonosítási aránya alacsony – ezt a szerzők korábbi vizsgálatok eredményeivel egyezőnek tartották, és további magyarázatra nem szoruló megfigyelésként jegyezték le. 2. Az izolált ejtésű magánhangzók azonosítási aránya magasabb a „nem rendes” ének módban, mint a „rendes”, operai módban. Ennek magyarázataként – bár arra egyértelmű akusztikai bizonyítékot nem találtak – a szerzők az $F_1 : f_0$ formánshangolás megjelenését feltételezték az operai ének módban. Magyarázatuk szerint tehát a kettes számmal jelölt eredményekkel kapcsolatban a két kondíció közti különbséget a hangolási stratégia megléte (a „rendes” operai produkciós módban) vagy hiánya (a „nem rendes” produkciós módban) okozhatta, hiszen a hangolás eltérést okoz a magánhangzónak a beszédben jellemző akusztikai szerkezetéhez képest. 3. A megemelt gégével képzett („nem rendes”) énekhangban az izolált ejtésű magánhangzók azonosítási aránya alacsonyabb volt, mint az azonos módban, de mássalhangzó-környezetben ejtett magánhangzók azonosítási aránya. A szerzők magyarázata szerint ennek legfőbb oka az lehetett, hogy a mássalhangzó-környezet a formánsátmenetek révén segítette a magánhangzók percepcióját még magas alapfrekvenciájú hangok esetében is. 4. Az operai technikával képzett énekhangban a mássalhangzó-környezetben ejtett magánhangzók azonosítási aránya magasabb volt, mint az azonos módban képzett énekhang izolált magánhangzói esetében, még magas alapfrekvenciákon is. A szerzők itt magyarázatként ismét csak azt sugallják (de óvatosan kerülnek a konklúzió explicit levonását), hogy ez az eredmény a mássalhangzó-környezetnek, valamint a formánsátmeneteknek köszönhető, tehát a mássalhangzó-környezet még a magas alapfrekvenciájú

énekelt magánhangzók esetében is segíti a magánhangzóknak az ejtési szándék szerinti azonosítását.

Fontos észrevennünk, hogy a szerzők utóbbi érveket annak ellenére javasolják, hogy a magasabb alapfrekvenciájú, operai ejtésű magánhangzók esetében maguk is formánshangolást, tehát a magánhangzónak az alapfrekvenciával változó akusztikai megvalósulásait feltételezték. Korábbi érvek értelmében a szerzők feltételezése szerint az $F_1 : f_0$ formánshangolás mint az operai ének mód sajátossága módosítja a beszédben jellemző akusztikai szerkezetet, és így csökkenti az ejtési szándék szerinti azonosíthatóságot. Ezzel magyarázható, hogy az izolált ejtésű magánhangzók azonosítása alacsonyabb volt az operai módban a „nem rendes” énekprodukciónak képest. Ha viszont ennek ellenére a mássalhangzó-környezetben mégis magasabb a magánhangzók ejtés szerinti azonosítása, mint izolált ejtésben, a mássalhangzók hatására vonatkozó magyarázat nem lehet a magánhangzók észleleti vetületére vonatkoztatható megfigyelés. A mássalhangzó-környezet pozitív hatásának a szerzők által javasolt feltételezése ugyanis azt jelenti, hogy a magánhangzókat az alapfrekvencia emelésével érintő akusztikai (és az e mögött meghúzódó artikulációs) változás, azaz az $F_1 : f_0$ hangolás a kontextus függvényében érvényesült. A negyedik pontban összefoglalt eredmény ugyanis a formánshangolás feltételezése mellett csak kétféleképpen értelmezhető: 1. a formánshangolás csak a kontextussal nem rendelkező magánhangzókat érintette, vagy 2. a formánshangolás a kontextussal rendelkező magánhangzókat az izolált magánhangzóknál kevésbé érintette. Amennyiben tehát a szerzők azt sugallják, hogy a mássalhangzó-környezet még a magas alapfrekvenciájú operai módban is segíti a magánhangzóknak az ejtési szándék szerinti azonosítását, az azt impikálja, hogy a szerzők eredményei és következtetései nincsenek teljes összhangban a kiinduló feltételezésekkel, hiszen érveik és magyarázataik ilyen módon nem a magánhangzók kontextusfüggő azonosítására, hanem azok kontextusfüggő megvalósítására (artikulációs és/vagy akusztikai sajátosságaira) vonatkoz(ná)nak.

Fontos megfigyelnünk azt is, hogy SMITH és SCOTT (1980) tanulmánya nem közöl sokat a percepció tesztek pontos körülményeiről – így nem tudjuk meg például, hogy az adatközlők milyen instrukciók alapján szolgáltatták válaszaikat. Annyi azonban bizonyos, hogy a válaszadás verbálisan történt – azaz a résztvevők szóban jelezték a kísérletvezetőnek, mely hangot vélték hallani –, valamint az is tudható, hogy a kísérletben használt CVC szekvenciák értelmes szavak voltak. A válaszadás modalitásából feltételezhetjük, hogy a kísérletvezető egyik kondícióban sem prezentált válaszlehetőségeket a résztvevőknek. Ha pedig nincsenek előre ismertett válaszlehetőségek, valószínűsíthető az is, hogy a kísérlet során más módon sem adtak megszorításokat a kutatók a hallgatóknak a válaszlehetőségek számának vagy minőségének tekintetében. Ha ez így van, az viszont azt jelenti, hogy míg a CVC kondícióban négy¹⁵ lehetőségből, négy értelmes szó közül választhattak a hallgatók, addig az izolált ejtésű

¹⁵ A választási lehetőségek száma elméletben ötre nő, ha feltételezzük, hogy az instruktor nem ismertette a szólistát a kísérlet elején, mert a fent listázottakon kívül szóba jöhet még a *bayed* szó is, azaz az /ei/ diftongus mint válaszlehetőség. Ugyanakkor az is feltehető, hogy a kitarított (azaz a monoftongusejtésnek kedvező) énekelt magánhangzók esetében ennek a hangzónak az előfordulása az észleletben nagyon alacsony valószínűségű,

hangok esetében a választási lehetőségek egyáltalán nem korlátozódtak arra a négy hangzóra, amelyet az értelmes szavak megengedtek. Az izolált ejtésben ugyanis a válaszlehetőségek számát csak az angol nyelv magánhangzókészlete korlátozta, mely legalább tíz¹⁶ lehetséges magánhangzóválaszt enged meg. Ennek fényében SMITH és SCOTT (1980) a magánhangzók-nak a mássalhangzó-környezet meglététől függő azonosítását valójában két jelentősen eltérő percepciós működést összevetve, két jelentősen eltérő kondícióban igyekezett vizsgálni. Az egyik kondícióban az értelmes szavak révén minden bizonnyal a nyelvi szerveződés magasabb szintjei, a morfológia, illetve a szemantika is befolyásolta az azonosítás tendenciáit, hiszen ezek korlátozták a válaszlehetőségeket és a válaszlehetőségek számát a válaszként szóba jöhető létező szavak (azaz a top-down percepciós folyamatok) révén. A másik, az izolált ejtésű hangzókat tartalmazó kondícióban ezzel szemben csak az alacsonyabb szintű, elemi hallási észlelésen és az akusztikai elemzésen alapuló (és csak az akusztikai kulcsokra támaszkodó, bottom-up) beszédpercepciós folyamatok léphettek működésbe úgy, hogy a lehetséges válaszok számát az angol nyelv magánhangzókészletén kívül semmi sem korlátozta. Nem állítható tehát bizonyosan, hogy a SMITH és SCOTT (1980) vizsgálatában kapott eredmények nem pusztán az egyes kondíciókban lehetséges válaszok számának, azaz az egyes válaszok előfordulási valószínűségének, valamint a helyes tippelési esély kondícióként eltérő mértékének következményei, hanem valóban a mássalhangzó-környezet pozitív hatását mutatják.

Végezetül pedig megemlítené az is, hogy a SMITH és SCOTT (1980) által vizsgált négy magánhangzó minősége is erősen befolyásolhatta a kapott eredményeket. A közelmúlt egyes, más kérdéseket megválaszolni kívánó, de a jelen problematika szempontjából érvényes eredményekkel szolgáló vizsgálatai alapján ugyanis ma már tudható, hogy a SMITH és SCOTT (1980) által vizsgált szavak és magánhangzók (egy hangzó kivételével)¹⁷ a gondozott, túlartikulált vagy hiperartikulált, lassú beszédben (idegen szóval a *clear speech*-ben) nagyobb számban ismerhetők fel, mint a hétköznapi, normál beszédben (ROGERS et al. 2010). Mivel pedig az éneklési feladatban SMITH és SCOTT vizsgálatában az énekes hosszú, azaz egyúttal feszesebb, „túlartikuláltabb” magánhangzókat ejtett, feltehető, hogy azon alapfrekvenciák esetében, amelyeken még az egyes magánhangzóknál nem lépett fel artikulációs-akusztikai változtatás (ti. $F_1 : f_0$ hangolás), az énekelte magánhangzóknak az ejtési szándék szerinti azonosítása a beszédhez hasonlóan jó arányú lehetett. Ráadásul azt is észre kell vennünk, hogy a SMITH és SCOTT (1980) által vizsgált négy beszédhangból kettő, mégpedig az /e/ és az /æ/ nyíltabb (harmadik és negyedik nyíltsági fokú) hangzók, melyekről nyíltsági fokuk és magas F_1 -értékük miatt feltehető, hogy még magas alapfrekvencián is jelentősebb változtatás

tehát ez a beszédhang jó eséllyel nem vagy csak kis számban merülne fel az ejtési szándéktól eltérő azonosítások esetén.

¹⁶ Csak az angol monoftongusokat számba véve ugyanis izolált ejtésben az /i ɪ e ε æ a ʌ o u ʊ/ hangzók jöhetnek szóba válaszként.

¹⁷ A kivételt a *lax*, azaz laza ejtésű /ɪ/ képezte, mely esetében a gondozott vagy túlartikulált beszédben a normál beszédejtéshez képest alacsonyabb volt a magánhangzók azonosításának aránya (ROGERS et al. 2010).

(azaz $F_1 : f_0$ hangolás) nélkül voltak ejthetők a kísérletben részt vevő énekes számára. Úgy tűnik tehát, hogy a magas alaphangfrekvencián is jelentősebb változtatás nélkül ejthető és a „túlartikuláció” miatt feltehetően jól azonosítható hangok SMITH és SCOTT (1980) vizsgálatában a „CVC” kondíció tesztanyagának pontosan felét képezték, ugyanakkor az egyes magánhangzók azonosítási tendenciáit csak összevonva ismertetik a szerzők. Ebből következően a magánhangzók minőségbeli jellemzőiből fakadó pozitív és negatív hatások összevegyülnek és kioltják egymást a vizsgálatban, ily módon valamelyest feltehetően félrevezető képet festve a tapasztalt észlelési tendenciákról.

Noha SMITH és SCOTT (1980) konklúziói alapján úgy tűnik, hogy a vizsgálat a mássalhangzó-környezet és a koartikulációs formánsátmenetek pozitív hatását mutatta ki a magas alaphangfrekvencián énekelt magánhangzók az ejtési szándék szerinti azonosítására, a fentebb ismertetett körülmények (a vizsgált magánhangzók és a kondíciók eltérő feltételeinek) fényében azt valószínűsíthetjük, hogy a vizsgálati dizájn *confound* (azaz több, egymással interakcióban lévő) hatás révén támogatta, erősítette a különbségeket a mássalhangzó-környezet hatását reprezentálni kívánó kondíció és az izolált ejtést reprezentáló kondíció között. Éppen ezért a szerzőknek a mássalhangzó-környezet pozitív hatására vonatkozó következtetését óvatosan kezelendőnek és további vizsgálatra érdemes kérdésnek tartjuk.

GOTTFRIED és CHEW (1986) az angol /i i e ε æ α ʌ o u u/ magánhangzókat vizsgálta férfi éneklésben, mégpedig egy kontratenor ejtésében. Észlelési tesztjeikkel a szerzők arra keresték a választ, hogy miként változik a magánhangzók azonosítása az f_0 és a fonációs mód vagy (zöng)regiszter függvényében (az utóbbi kérdést a mondális zöngé és az általában a férfiak által a magasabb alaphangfrekvenciákra használt falzett szembenállásával reprezentálva), valamint a mássalhangzó-környezet hatására. Ehhez a *bVd* hangkörnyezetben (azaz tíz i. létező angol szóban, ii. létező szavak egy-egy önmagában értelmetlen szótagjában, illetve iii. értelmes szavakra rímelő álszóban) rögzítették az imént felsorolt magánhangzókat mindkét fonációs módban: modális zöngével az *c* (130 Hz), *e* (165 Hz), *a* (220 Hz), *c'* (260 Hz) és *e'* (330 Hz), falzett fonációval pedig az *a* (220 Hz), *c'* (260 Hz), *e'* (330 Hz), *a'* (440 Hz), *c''* (520 Hz) zenei hangokon (a kétféle fonációs móddal produkált alaphangfrekvencia-tartományok tehát átfedésben voltak). Ezután a kutatók kivágták minden szótag középső, megközelítőleg 200 ms hosszúságú kváziszacioner részét (azaz a magánhangzó „tisztá fázisát”, vö. GÓSY 2004), így létrehozva egy, a mássalhangzó-környezetben ejtett magánhangzókkal, azaz a „teljes szótag” kondícióval opponáló „magánhangzó” kondíciót. A vizsgálat során a szerzők összesen négy percepciós tesztet végeztek el. Ezek során rendre a következő feltételpárokat tesztelték: „modális” + „teljes szótag”, „modális” + „magánhangzó”, „falzett” + „teljes szótag”, „falzett” + „magánhangzó”. Az egyes opponáló feltételeket („teljes szótag” és „magánhangzó”, valamint „modális” és „falzett”) tehát külön-külön vizsgálták.¹⁸

¹⁸ A mássalhangzó-környezetben ejtett magánhangzók azonosításához a kutatók a tesztek során kitéltetett feladatlapokon minden esetben értelmes szavakat jelentettek meg hordozó hangsorként, melyek vagy a teljes szótagnak, vagy az önmagukban értelmetlen szótagokhoz tartozó értelmes szavaknak, vagy pedig a hallott álszókra rímelő valódi szavaknak feleltek meg: *bead* 'gyöngy', *bid* 'ajánlat, tét', *bade* 'kínál' (múlt idő), *bed*

A tanulmányban ismertetett eredmények szerint mind az f_0 , mind pedig a fonációs mód hatással volt a magánhangzók ejtési szándék szerinti azonosításának százalékos arányára. Az ejtési szándéktól eltérő válaszok aránya az f_0 emelésével nőtt, ám a modális regiszter felső tartományában talált nagy „hibaszázalékot” a fonációs mód váltása, azaz a falzett képzésmódra való áttérés csökkentette, míg az f_0 további emelésével az ejtési szándéktól eltérő válaszok aránya ismét nőni kezdett. Ezen túlmenően a vizsgálat azt is kimutatta, hogy a veláris hangzókra átlagosan nagyobb arányban érkezett az ejtési szándéktól eltérő válasz, valamint azt is, hogy a teljes szótagra minden alapfrekvencián nagyobb arányban érkeztek az ejtési szándéknak megfelelő válaszok, mint a magánhangzók tiszta fázisát tartalmazó stimulusokra. Az akusztikai elemzések (LPC-analízis, melynek eredményei a magas alapfrekvencián tapasztalható spektrális alul-mintavételezés mint megszorítás figyelembevételével értelmezendők) szerint a regiszterek, azaz a kétféle zöngképzési mód közti különbségek az F_1 és F_2 formáns értékére is hatással voltak, mégpedig a következőképpen. A beszédben alacsonyabb F_1 formánssal rendelkező első, második vagy harmadik nyíltsági fokú hangzók esetében a falzett módban a beszédértékhez közelebbi F_1 volt mérhető, míg ugyanezen fonációs módban a zárt és veláris hangzókban a beszédben mérhetőhöz közelebbi frekvencián megvalósuló F_2 -t találtak. (Megjegyzendő, hogy nem teljesen egyértelmű, hogy a szerzők az alacsonyabb F_1 és F_2 értékeket nem az LPC-analízis téves formánsdetekciójaként értelmezik-e.) GOTTFRIED és CHEW (1986) következtetései szerint a magasabb alapfrekvencián ejtett magánhangzókra kapott alacsonyabb azonosítási arányok oka egyaránt (és akár egymástól függetlenül) lehet a magánhangzók az alacsonyabb alapfrekvenciákon tapasztalttól eltérő artikulációja, valamint a magas alapfrekvencián a spektrumban megjelenő felharmonikusok kisebb száma is (azaz a magánhangzók alacsony felbontása vagy másként a magas alapfrekvencián ejtett szonoránsokat jellemző spektrális alul-mintavételezettség, vö. 1.1.1. *alfejezet*).

Tanulmányuk bevezetőjében GOTTFRIED és CHEW (1986) hivatkozik azokra az egyes hangszerek hangjának felismerését vizsgáló korábbi pszichoakusztikai eredményekre, melyek szerint az egy adott alapfrekvencián megszólaló hang hangszínének azonosításában, azaz a hangszerek keltette hangok világában a hangot kiadó hangszer felismerésében kiemelt szerepű akusztikai kulcs a hangindítás; az idézett tanulmányok eredményei szerint ugyanis a hangindítás eliminálásával jelentős mértékben csökkent a hangszerek felismerésének sikeressége (pl. BERGER 1964, SALDANHA–CORSO 1964). Bár az említett eredmények citálása azt a benyomást kelti, hogy a szerzők elfogadják azt a feltételezést, mely szerint a hangszínekben eltérő hangok megkülönböztetésében a hangadás kezdetének kiemelt jelentősége van, valamint úgy tűnik, hogy a feltételezést a beszédhangokra (tehát két különböző minőségű magánhangzó megkülönböztetésére nézve) is kiterjeszhetőnek tartják, érdekes – és egyben

¹ágy', *bad* 'rossz', *bod(y)* 'test', *bode* 'jósol, előre jelez', *bud* 'pajtás', *bood* (*good*) 'jó', *boood* (*food*) 'étel'. A másalhangzó-környezetből kivágott magánhangzókhoz a szerzők szintén egy szótagos értelmes szavakat prezentáltak válaszlehetőségek gyanánt, mely szótagok (értelemszerűen) az adott magánhangzókot tartalmazták: *eat* 'eszik', *it* 'az', *ate* 'eszik' (múlt idő), *Ed* (férfinév), *at* '-nál', *odd* 'páratlan', *up* 'fel', *oat* 'zab', *hook* 'kampó', *ooze* 'izzad' (GOTTFRIED–CHEW 1986: 125).

igen sajnálatos – módon mégsem ellenőrzik azt vizsgálatukban. Az izoláltan ejtett hangok helyett ugyanis csak a szótagmagból kivágott kvázistacioner magánhangzószeleteket tesztelik. Mivel pedig ily módon a tesztben izoláltan lejátszott magánhangzók nem valódi izoláltan ejtett magánhangzók voltak (melyek ennél fogva az izolált magánhangzók karakterisztikus hangindítási sajátosságaival sem rendelkeznek), a vizsgálat eredményei nem vonatkoztathatóak egyértelműen a valódi izoláltan ejtett, illetve a mássalhangzó-környezetben ejtett magánhangzók azonosításának különbségeire. Ezt az egyébként releváns és érdekesnek ígérkező összehasonlítást tehát GOTTFRIED és CHEW (1986), bár implicite felvetni látszik, mégis nyitott kérdésként hagyja.

GOTTFRIED és CHEW (1986) egyik javaslata szerint tehát a magánhangzóknak a kísérletben tapasztalt, az f_0 emelésével összefüggően egyre bizonytalanabb (az ejtési szándéktól eltérően kategorizálható) hangszíne nem feltétlenül az f_0 emelését kísérő artikulációs és az azzal járó akusztikai változások, hanem a spektrális alul-mintavételezettség következménye is lehet. Amennyiben pedig ez a feltételezés igaz, és a magánhangzókat tényleg nem jellemzi más – jelentősebb – változás, mint a fokozódó spektrális alul-mintavételezettség, jól beláthatónak tűnik a szerzők azon magyarázata is, mely szerint (a kísérletben vizsgált 520 Hz alatt) a mássalhangzó-környezet a koartikulációs formánsátmenetek révén segítheti az ejteni kívánt magánhangzók azonosítását. Ha azonban a magánhangzók artikulációs és akusztikai szerkezetében az f_0 -tól függő változásokat, tehát például $F_1 : f_0$ hangolást is feltételezünk, amint arra a szerzők második elképzelhető magyarázatként utalnak (és amire gondolni a már ismertett akusztikai elemzések alapján olvasóként is jó okunk lehet), ez a magyarázat nem állja meg a helyét. Ebben az esetben ugyanis a módosult artikulációval ejtett magánhangzók és a mássalhangzók közti átmenetek logikusan csak a módosult képzésű magánhangzót kulcsolhatnák a percepció számára, nem pedig az ejteni kívánt (de valójában nem megvalósított) magánhangzót. Ha tehát az ejteni szándékozott magánhangzók artikulációs-akusztikai megvalósítása módosul az alapfrekvencia emelésének következtében, akkor a mássalhangzó-környezettől sem várható, hogy az ejteni szándékozott minőségre utaljon, a kontextus legfeljebb csak a módosult akusztikai szerkezet pontosabb felismerését segíthetné. A szerzők ez utóbbi magyarázatot azonban nem fejtik ki részletesen, így nem tudjuk meg például azt sem, hogy mivel tartják magyarázhatónak azt az elemzésükkor alkalmazott elvet, mely szerint az esetlegesen megváltozott artikulációval ejtett magánhangzókra adandó „helyes válaszként” egyértelműen az íráskép által megjelölt magánhangzót feltételezik. A tanulmány nyomán további kérdésként merül fel az is, hogy mit tapasztalának a szerzők az általuk vizsgált kondíciókban az igazán magas alapfrekvenciákon – az ismertett vizsgálat ugyanis csak egy viszonylag alacsony alapfrekvenciáig (körülbelül 500 Hz-ig) fedi le a magas hangosztályokba tartozó énekesek által használt frekvenciatartományt (mely a klasszikus értelmezésben 1047 Hz-ig, azaz a „magas céig” terjed).

SCOTTO DI CARLO és GERMAIN (1985) 15 francia magánhangzót elemzett egy percepció vizsgálatban egy koloratúrszoprán izolált ejtésében, az énekes teljes alaphangtartományán

rögzítve azokat 220 Hz és 1046 Hz között 17 különböző alaphfrekvencián. A vizsgálatban a szerzők az /i e ε a u o ɔ/ orális és az /ẽ œ õ ã/ nazooralis monoftongusokat használták fel, és a magánhangzókat izolációban rögzítették, majd izolációban játszották le az észlelési tesztben háromféle randomizált sorrendben. A percepció vizsgálatban négy fonetikailag képzett, de az éneklésben (és zenében) nem jártas adatközlő vett részt. Az instrukciók szerint az adatközlők feladata az volt, hogy azonosítsák, majd fonetikai szimbólumokkal jelöljék az elhangzott magánhangzókat. A tesztek megelőzte egy rövid familiarizációs szakasz is, mely során a kísérletben részt vevő négy adatközlő meghallgatott egy, az énekes által énekelt rövid operarészletet (felvételtől), melyhez a librettót is elolvashatták. Ezt a „tanítási fázist” (mely során a résztvevők valamelyest megismerkedhettek a stimulusokat szolgáltató énekes hangjával, hangképzésével és énekelt magánhangzóinak sajátosságaival) a rövidtávú memória kondicionálásának tekintve a szerzők úgy vélték, hogy ez által módjuk volt egységesíteni és következetesebbé tenni az egyébként az egyes stimulusokra adott, akár egyénenként is igen változatos megoszlást mutató válaszokat. Kísérletében SCOTTO DI CARLO és GERMAIN elsősorban az alaphfrekvencia hatását vizsgálta, de külön kondícióként a regiszterek hatását is elemezték. E faktor kezelésében SCOTTO DI CARLO és GERMAIN (1985) tanulmánya jól reprezentálja a szakirodalomban egyébként is honoló nagy terminológiai és fogalmi bizonytalanságot. A szerzők ugyanis ötféle regisztert különböztettek meg mint alaphfrekvencia-tartományt, rendre alsó (*lower*), alsó-közép (*lower middle*), középső (*middle*), felső-közép (*upper middle*), és felső (*upper*) regisztert, melyekhez produkciós vagy zöngképzési módokat rendeltek, de élesen elkülönítve azokat az előbb felsorolt, az alaphfrekvencia-tartományok megnevezését szolgáló fogalmaktól. A kísérletben elkülönített zöngképzési módokat (nem a ma korszerűnek elfogadott terminusokkal élve) a szerzők rendre mellkasi (*chest voice*), alsó közép- (*chest mid voice*), közép- (*mid voice*) és fejhangnak (*head voice*) nevezték.¹⁹ Vélekedésük szerint bár az utóbbiak valóban az előbbiekhöz rendelhetők, ez a hozzárendelés nem

¹⁹ A tanulmányban használt terminológia korszerűtlenségét éppen az mutatja, hogy bár a szerzők állítása szerint a fogalmak fonációs módokat jelölnek, a megnevezéshez mégis az adott módban történő hangadást kísérő szenzoros rezonanciaélmények leírását (ti. pl. *chest voice* = „rezgés a mellkasban”) használják. Mivel ma már a regisztereket egyértelműen gége- és hangszalagszintű, kizárólag a fonációhoz kötődő jelenségeknek tartja a szakirodalom, ezért ezeket a korábban elterjedt megnevezéseket félrevezető mivoltuk miatt egyértelműen kerülni igyekszik. A fonációs módoknak a kísérletben definiált homályos voltához járul még, hogy az itt *mellkasi hang* és *középhang* terminusokkal megnevezett és egymástól elkülönített produkciós módoknak nem feleltethetők meg egyértelműen a zöngképzés módozataival foglalkozó kurrens szakirodalomban leírt produkciós (azaz zöngképzési vagy rezgés-) módok. A jelenleg legáltalánosabban használt modellek ugyanis (a használati alaphfrekvenciák szerint rendre) a következő regisztereket különböztetik el: a) irreguláris zöngé vagy glottalizáció (*vocal fry*, *creaky voice*, *pulse*, 0. mechanizmus vagy M0), b) a mindennapokban leggyakrabban használt modális zöngé (*modal*, *chest*, *voix mixte [for men]*, 1. mechanizmus vagy M1), c) falzett (*falsetto*, *head*, *voix mixte [for women]*, *mixed*, 3. mechanizmus vagy M3), valamint d) füttyregiszter (*falgolet*, *flute*, 4. mechanizmus vagy M4). Ezeket tekinti tehát a szakirodalom rezgésmódjában eltérő zöngképzési alaptípusoknak (vö. HENRICH et al. 2004; TITZE 2004; ROUBEAU et al. 2009; SEIKEL et al. 2010; MARKÓ 2013). A teljesség kedvéért e kitérőben megjegyezzük azt is, hogy bár a fent felsorolt alaptípusokhoz képest a sokat vizsgált *pressed* vagy „préselt” és *breathy* vagy leheletes zöngképzések hangszínezetbeli eltéréseket mutatnak, ezek nem az előbbiektől eltérő vibrációs mód révén jönnek létre, így azok a besorolások, melyek a fonációs módot mint hangszalag-vibrációs mechanizmust tekintik, ezeket nem különböztetik el zöngképzési alaptípusként (vö. pl. SEIKEL et al. 2010).

egyértelmű, hiszen egyes alaphangtartományok egyes hangjai többféle képzési móddal is képezhetők lehetnek. Épp ebből kifolyólag azonban nehezen értelmezhető, hogy mégis miért, és főként hogy miként használhatták összehasonlításaikban és a statisztikai elemzésekben a produkciós módokat az alaphangtartomány helyett, hiszen arról, hogy a zöngéképzést vagy hangszalagrezgést valamely endoszkópián vagy más elven alapuló eszközzel megfigyelték volna, nem számolnak be a szerzők. Feltehető, hogy a fogalmi szétválasztás ellenére a szerzők végül a „produkciós módok” megnevezésen mindvégig az alaphangtartományokat (és azok tartományait) értik. Mivel így a tanulmányban a „regiszterek” és „fonációs módok” fogalmi folyamatosan keverednek, és ebből kifolyólag a szerzők számos következtetése bizonytalan (vagy azért, mert a zöngéképzésből következő, de a frekvenciatartománnyal kapcsolatban értendő, vagy fordítva), az egyes „regiszterekre” vagy „fonációs módokra” kapott eredményeket a jelen értekezésben nem tárgyaljuk. Sorra vesszük most azonban az általuk vizsgált további paraméterekre kapott eredményeket.

SCOTTO DI CARLO és GERMAIN (1985) eredményei szerint a magánhangzók az ejtési szándék szerinti azonosítása csökkent az alaphangtartomány emelésével, mely tendenciára mind a magánhangzó-minőség, mind pedig a produkciós mód (vagy talán inkább az alaphangtartomány), mind pedig a kettő kombinációja befolyással bírt. Mivel azonban sajnos a tanulmányban a legtöbb esetben az egyes „regiszterekre”, azaz alaphangtartományokra összevont adatokat közölnek a szerzők, így sok esetben nem tudjuk pontosan meghatározni, mely alaphangtartományokhoz köthető egy-egy megfigyelt észlelési tendencia megjelenése. A több alaphangtartomány adatait egyetlen adatpontba összesítő elemzésekből és ábrázolásokból azt látjuk, hogy a vizsgált magánhangzók az ejtési szándék szerinti azonosítása az f_0 emelésével megközelítőleg fokozatosan csökkent, de olyan módon, hogy egy ponton (a közölt táblázatukból kikövetkeztetve megközelítőleg 349 és 392 Hz, azaz az f' és a g' zenei hangok között) az azonosítási százalék drasztikus esést mutatott (62%-ról 38%-ra), és végül a legmagasabb alaphangtartományon 9%-ot ért el.

SCOTTO DI CARLO és GERMAIN (1985) az azonosítási tendenciákat a magánhangzók egyes distinkatív fonológiai jegyei, mégpedig az ajakműködés (réses és ajakkerekítéses magánhangzók szétválasztása), az állkapocsnyitás (nyílt és zárt magánhangzók szétválasztása), a nazoorális minőség (nazoorális és orális magánhangzók szétválasztása), valamint a képzési hely, a toldalékcsovében fellépő szükület helye és a nyelvállás szerint is elemezték. Ez utóbbi szempont, azaz a nyelvállás és az az szerinti besorolás azonban végül tisztázatlan marad, ugyanis erről sem a módszertani bevezetőben, sem pedig másutt nem tudunk meg pontosabban a tanulmány során.

Az eredmények szerint a fonológiai jegyeknek az ejtési szándékhoz megfelelő azonosítási arányai a magánhangzó-minőségek általános azonosítási tendenciáihoz hasonlóan az alaphangtartomány emelésével fokozatosan csökkentek. A vizsgált jegyek között a labialitás tűnt a legkevésbé ellenállónak, ugyanis ennek az ejtési szándékhoz megfelelő azonosítása a legmagasabb alaphangtartományon (a 988 Hz-re és az 1047 Hz-re kapott eredmények összevonásával kalkulálva) 1,5%-ig csökkent. Ezt a szerzők az általuk feltehetőleg a felvételek közben

megfigyelt és így (szintén általuk) a magasabb alapfrekvenciákon általánosan jellemzőnek tartott ajakréses ejtéssel magyarázták, és magyarázataikat az /i/-re és /e/-re magas alapfrekvenciákon talált relatíve magas, az ejtési szándék szerinti azonosítási arányokkal is alátámasztottnak vélték.

Bár nem derül ki egyértelműen, hogy mely magánhangzókat értik a szerzők „zárt” és „nyílt” magánhangzókon (nem használnak ugyanis a köztes fokozatokra vonatkoztatható megnevezéseket, tehát nem egyértelmű az /e ε a ø œ o ɔ/ orális valamint az /ē ē̃ ē̄/ nazoorális monoftongusok besorolása), a szerzők egy további eredménye szerint a „zárt” magánhangzóknak az ejtési szándék szerinti azonosítása az alacsonyabb alapfrekvenciákon magas volt (46% körül), az f_0 emelésével pedig egyre csökkent (a legmagasabb alapfrekvencián tapasztalt 4%-ig). Ezzel szemben azonban a „nyílt” magánhangzóknak az ejtési szándék szerinti azonosítása alacsonyabb alapfrekvenciákon sem volt túl magas (maximum 46%, amit a legalacsonyabb alapfrekvencián tapasztaltak), a magasabb alapfrekvenciákon pedig a zárt magánhangzókhoz hasonlóan a nyíltakra kapott ejtés szerinti azonosítás aránya is csökkent (bár itt csak 19%-ig). Ezeket az eredményeket a szerzők a következőképp magyarázzák. Amint azt korábbi produkciós vizsgálatok bizonyították, a magas alapfrekvencián ejtett magánhangzók artikulációsan nagyobb állkapocsnyitásszöggel valósulnak meg. Éppen ezért – folytatják a szerzők – az is jól belátható, hogy miért nehezebb magas alapfrekvencián a zárt hangzók azonosítása. Nem térnek ki azonban annak az ellentmondásnak a feloldására, mely e között és az előző következtetésük között feszül: nem értelmezik ugyanis, hogy a magas alapfrekvenciákon alkalmazott nagyobb állkapocsnyitás ellenére miként lehetséges mégis az, hogy az /i/ és /e/ még magas alapfrekvencián is „relatíve jó arányban” volt az ejtési szándék szerint azonosítható. (Táblázatok tanúsága alapján egyébként ez az arány százalékokban kifejezve az f' [698 Hz], g' [784 Hz], a' [880 Hz], h'' [988 Hz] és c''' [1046 Hz] hangokon rendre 50, 25, 16, 25 és 16% az /i/ esetben, valamint 50, 8, 16, 16 és 0% az /e/ esetben). Még jobban megnehezíti az eredmények értelmezését, hogy a szerzők érveikben értelemzavaróan egybemossák a produkciós és percepciós tényezőket. A magasabb alapfrekvenciákon ugyanis, a korábbi vizsgálatok szerint, nem a zárt magánhangzók azonosítása nehéz, hanem a hangzók ejtése lehetetlenül el. Hiszen, amint azt a 1.2.2. *alfejezet*ben is láttuk, az énekes magas alapfrekvenciákon énekelve és az f_0 további emeléséhez az F_1 értékét az f_0 -ra (vagy a fölé) hangolja, amit az állkapocsnyitásszög növelésének (vagy a nyelvállásfok csökkentésének) segítségével ér el. Ha tehát már eleve nyíltabb hangzókat ejt az énekes, nem a zárt hangzók egyre nehezülő „felismeréséről” kell beszélnünk (a szerzők az *identification* szót használják, vö. SCOTTO DI CARLO–GERMAIN 1985: 194), hanem legfeljebb a zárt magánhangzó-minőség megjelenéséről az észleletben, hiszen már produkciósan sem (egyértelműen) zárt az adott hangzók minősége.

A nazális jeggyel kapcsolatos (elégge elnagyolt, az összes alapfrekvenciát összesítő) összehasonlításban a szerzők azt találták, hogy az ejtési szándék szerint nazoorális magánhangzók nazoorálisiként való azonosítása alacsony volt. Erre magyarázatképp a következő feltételezést fogalmazzák meg: mivel a teljesen nazalizált nazoorális hangszín az éneklésben nem kívánatos, ezért az énekes igyekezett kerülni azt, mégpedig úgy, hogy lehetőség szerint csak igen

kismértékben (vagy ha ez nem lehetséges, inkább egyáltalán nem) nazalizálta a magánhangzókat. (A szerzők szerint ugyanis az énekelve ejtett nazooralis magánhangzók ejtések az elsődleges artikulációs cél az orális magánhangzó, melyet az énekesek csak utóbb nazalizálnak – vagy nem nazalizálnak olyan esetekben, amikor a kismértékű nazalizáció megvalósítása aerodinamikai vagy más okokból nehézségekbe ütközne.)

Végezetül a szerzők a „hibaként” legnagyobb számban megjelenő magánhangzók rangsorát elemzik. Ez alapján megállapítják, hogy a tesztelők az ejteni szándékozott hangzók helyett legnagyobb arányban az /a/ hangzót jelölték meg (mégpedig az összes „hiba”, azaz az ejtési szándéktól eltérő azonosítás 34%-ában). Ezt az eredményt a nyíltság esetében megfogalmazott magyarázatuk ismétlésével értelmezik, azaz utalnak rá, hogy a magas alapfrekvenciákon produkált énekhang az /a/ ejtésének kedvez, valamint arra is, hogy a nyíltabb artikuláció vélekedésük szerint egyben nyíltabb észleletet is jelent, aminek értelmében a nagyszámú /a/ válasz a szerzők szerint a különféle magánhangzók az énekes által egységesen /a/-ként történt megvalósítására utal. Mivel azonban ez a lista az összes alapfrekvencián az összes magánhangzóra kapott válaszokat tartalmazza, nem tudjuk meg pontosan, hogy specifikusan mely magánhangzók esetében mely magánhangzóválasz volt jellemző.

SCOTTO DI CARLO és GERMAIN (1985) magyarázatai tehát egyértelműen azon alapszanak, hogy a korábban artikulációs vizsgálatok során demonstrált produkciós tendenciákat közvetlen, lineáris összefüggésbe hozzák a vizsgálatokban talált percepciós adatokkal – így például az artikulációs nyíltság számukra könnyen beláthatóvá teszi a percepciós nyíltság megjelenését. Eredményeik értelmezésekor azonban komoly problémát jelent, hogy hiába közlik a szerzők a „zárt” magánhangzók „nyíltabbként” történt azonosításának tendenciáját, nem tudjuk meg e „zártabb” magánhangzók pontos minőségét. Így nem tudjuk megítélni például azt sem, hogy állításukat olyan magánhangzókra alapozzák-e a szerzők, melyeket (a magánhangzórendszerből következően) egyaránt lehetne nyíltabbként vagy zártabbként is azonosítani, avagy olyan hangzókra (is), melyeket – a rendszer legzártabb hangzói lévén – az ejtési szándéktól eltérő azonosítás esetén szükségszerűen eleve csak nyíltabbként lehet felismerni.

Ráadásul az akusztikum és a percepció közti egyszerű lineáris összefüggéseknek a feltételezése sokszor ellentmondásokhoz is vezet a szerzők érvei között. Példaként említhető, hogy a szerzők figyelmen kívül hagyva az egyes fonológiai jegyek azonosítása közötti összefüggéseket következtetéseik levonásakor az /i/ és /e/ magas alapfrekvencián magasabb arányú az ejtési szándékkal egyező azonosítását egyszer (az ajakműködés szerint) indokoltnak tartják, miközben máskor, egy másik érvelésük értelmében (a nyíltság szerint) ugyanez a magas arányú azonosítás indokolatlan lenne (hiszen a szerzők azt feltételezik, hogy a nyíltabban ejtett hangzókat nyíltabbként is azonosítják a hallgatók).

Ezen túlmenően fontos megfigyelnünk azt is, hogy a szerzők nem tesznek említést formánshangolásról, sem pedig az akusztikai szerkezetet az alapfrekvencia függvényében érintő más akusztikai változásról, ugyanis az észlelésre kapott eredményeiket közvetlenül és kizárólag a (korábbi leírásokban közölt) artikulációs megfigyelésekkel hozzák összefüggésbe.

Ilyen értelemben a vizsgálatukban alkalmazott értelmezési keret talán leginkább a motoros elmélet (vö. *I.1.2.1. alfejezet*) valamifajta változatának lenne tekinthető, de mint ilyen is nagyon leegyszerűsítő és egyúttal sajnos szükségszerűen félrevezető is. Mivel tudjuk, hogy mind az artikuláció és az akusztikum, mind pedig az akusztikum és percepció viszonyát bonyolult összefüggések jellemzik, SCOTTO DI CARLO és GERMAIN (1985) következtetései csak nagymértékű körültekintéssel, óvatosan kezelendők.

Bár a fent felsorolt okokból úgy véljük, hogy SCOTTO DI CARLO és GERMAIN (1985) vizsgálatából nem vonhatunk le jól megalapozottnak tekinthető és jól általánosítható konklúziókat, mégis úgy látjuk, hogy a kísérlet számos izgalmas kérdést vet fel, melyeknek ellenőrzése kísérletes vizsgálatot igényel. A szerzők adatai ugyanis rámutatnak például arra, hogy az alapfrekvencia emelésével nem feltétlenül csak a legnyíltabb magánhangzók észlelete maradhat relatíve magas arányú (lásd az /i/-t és /e/-t), tehát gyümölcsöző lehet a magánhangzó-minőségek minél szélesebb körű összehasonlítása. Ezen túlmenően pedig nem szabad megfeledkeznünk arról sem, hogy ez a vizsgálat tekinthető az első olyan percepció-kísérletnek is, melynek céljai között szerepelt az egyes különböző nyíltsági fokú magánhangzók szétválasztása az észlelési vizsgálatban, ami a korábbi artikulációs-akusztikai eredmények alapján erősen motivált. Így ez a vizsgálat mindenképpen fontos mérföldkő az énekelt magánhangzók észlelésére vonatkozó kutatások történetében.

BENOLKEN és SWANSON (1990) az angol /i i ε u o ɔ/ monoftongusokat, valamint az /ei/ diftongust vizsgálták egy szoprán éjtésében a *hVd* hangkörnyezetben (tehát hat értelmes angol szóban), 13 különböző alapfrekvencián 262 Hz alapfrekvenciától felfelé, két oktáv alaphangterjedelemben. Az így rögzített hanganyagot 28 résztvevő segítségével vizsgálták egy percepció-kísérletben. Az adatközlők egy teremben ülve, hangszórókon keresztül hallgatták az egybefüggő felvételt, és az elhangzott *hVd* hangsorok magánhangzóinak azonosított hangminőségét papíron rögzítették. A szerzők azt találták, hogy az alapfrekvencia emelésével egyre csökkent az éjtési szándéknak megfelelő magánhangzó-minőségek azonosításának sikeressége. Ezzel együtt az f_0 emelésével a zárt, első nyíltsági fokú /i/ hangot egyre nagyobb számban azonosították az adatközlők az ejteni szándékozott hangzó minőségénél nyíltabb /i/-ként majd /ε/-ként, valamint a zárt, első nyíltsági fokú /u/ hangot is egyre nyíltabb, azaz /o/ majd /ɔ/ magánhangzókként azonosították a tesztben részt vevők. Ezen felül a szerzők egy másik összefüggésről is beszámolnak. Azt találták ugyanis, hogy a nagyon magas, 1047 Hz alaphangfrekvenciájú zenei alaphang magasságában a palatális és veláris hangzók csoportjára két, egymástól eltérő spektrális kép, illetve észlelési mintázat volt jellemző. A spektrum tekintetében a velárisok csoportjára az első (f_0) és a harmadik ($3f_0$) harmonikus nagyobb, míg a második ($2f_0$) és negyedik ($4f_0$) harmonikus kisebb intenzitása volt jellemző – erre az észleletben főként veláris hangok, különösen az /ɔ/ és /a/, valamint a palatális /æ/ jelentek meg válaszként. Ezzel szemben a palatálisok esetében az első harmonikustól a negyedikig fokozatosan csökkenő amplitúdó volt jellemző – erre az észleletben (a válaszok között) veláris és palatális hangzók egyaránt megjelentek (a válaszok megoszlásában a velárisoknál tapasztalt szórásnál nagyobb szóródást mutatva).

A szerzők értelmezése szerint a beszédhangok az ejtési szándéknál nyíltabb hangokként való azonosítása a korábban leírt $F_1 : f_0$ hangolási tendenciákkal egybevágó eredmény, vélekedésük szerint tehát az akusztikum és a percepció összefüggése a magas alapfrekvenciájú énekelt magánhangzók esetében lineáris, ahol a nyíltabb akusztikai vetület nyíltabb percepció vetületet is implikál. A magas alapfrekvenciákon talált kétféle spektrális mintázatból ezen túlmenően pedig arra is következtettek, hogy az 1047 Hz alapfrekvencián 500 Hz alatt megtalálható négy harmonikus ($f_0, 2f_0, 3f_0, 4f_0$, tehát az alaphang és az első három felharmonikus) elegendő információt hordoz(hat) ahhoz, hogy a palatális és veláris magánhangzócsoportok még ilyen magas alaphangon is (bizonyos mértékig) elkülöníthetők maradjanak. Sajnos azonban a szerzők adósok maradtak azzal a nem magától értetődő magyarázattal, hogy milyen módon volna párhuzamba hozható a magas alapfrekvencián talált spektrális mintázat az alacsonyabb frekvenciákon jellemző formánsszerkezettel. Nem adnak tehát választ (vagy fogalmaznak meg javaslatot) arra a kérdésre, hogy vajon miként feleltetheti meg a beszédpercepció a magasabb alapfrekvenciákon megvalósult énekelt magánhangzók (alul-mintavételezett és a formánshangolás miatt megváltozott) spektrális szerkezetét az alacsonyabb alapfrekvenciákon tipikusan megszokott spektrális szerkezeteknek.

HOLLIN és munkatársai (2000) az egymástól artikulációs és akusztikai tekintetben legmesszebb elhelyezkedő (amerikai angol) /i u a/ magánhangzók észlelését vizsgálták 13 haladó technikájú nő és 5 férfi énekes ejtésében „magas” és „mély” alaphangokon, „halk” és „hangos” hangadásban. A nők esetében a „mély” alaphangok a *G* (98 Hz) vagy az *A* (110 Hz) zenei hangokat jelentették, a „magas” alaphang esetében pedig mindössze annyit fogalmaztak meg kérésként a kutatók a felvételek készítésekor az énekeseknek, hogy legalább egy *a*''-t (880 Hz), de lehetőség szerint még magasabb hangot produkáljanak. (Az énekesek tehát a megjelölt hangokból tetszőlegesen választhattak, így az alapfrekvencia az egyes énekesek ejtésében a „mély” és „magas” kondíciókon belül nem feltétlenül egyezett). Az eredmények szerint a „mély” kondícióban a női énekesek által elért alapfrekvencia átlagosan 200 Hz körül, a „magas” kondícióban mérhető átlagos alapfrekvencia pedig 900 és 1000 Hz között mozgott. A percepció teszthez a kutatók a magánhangzók közepéből (a magánhangzók tiszta fázisának tekinthető) 1,5 másodperces szeleteket vágtak ki, majd a hangzókat randomizált sorrendben játszották le a 38 képzett (beszédtanár vagy fonetikus végzettségű, illetve beszédterapeuta-hallgató) és 12 nem képzett résztvevőnek. A percepció vizsgálatot akusztikai elemzés is kísérte, a szerzők az énekelt magánhangzók első két formánsának frekvenciaértékét mérték (hagyományos, spektrális elemzés segítségével).

A percepció teszt eredményei szerint bár a hangerősség bizonyos mértékben eltérően hatott az egyes hangokra (az /i/ és /a/ nagyobb hangerő mellett inkább az ejtési szándék szerint volt felismerhető, míg az /u/-ra a hangerősség éppen fordított hatást gyakorolt), ez a különbség (statisztikailag) nem volt számottevő. Az ejtési szándék szerinti azonosítás (a hangerősség figyelembevétele nélkül átlagolva) az /i/ esetében az alacsonyabb alapfrekvencián 26,7%, majd a magasabbon 2,2% volt, az /u/ esetében 38%, majd 1,8%, az /a/ esetében pedig 74,2, majd 45,5%. Ezek alapján a gyakorlatilag minden alapfrekvencián az ejtési szán-

déktól nagy arányban eltérő azonosítást mutató eredmények alapján (melyekben az f_0 emelésével csökkenés is látható volt) a szerzők azt a következtetést vonták le, hogy adataik a korábbi eredményekkel egybevágnak. Mivel azonban az akusztikai adatok alapján az alacsonyabb alapfrekvenciákon még nem lépett fel $F_1 : f_0$ hangolás, az ezeken az alapfrekvenciákon talált alacsony azonosítási arányokat a szerzők nem tudták megmagyarázni. Ezen túlmenően, a magas alapfrekvenciákon talált, az ejtési szándéktól eltérő azonosítási tendenciákkal kapcsolatban a szerzők azt is kiemelik, hogy azokat a beszédben magasabb F_1 -gyel rendelkező /a/ válaszok dominálják (44-48% körüli válaszaránnyal). Ugyanakkor táblázataikból az is jól látható, hogy ezeken az alapfrekvenciákon a fonológiai laza, zárt képzésű /u/ válaszok is nagyobb (15-16% körüli) arányban jelentek meg mind a három ejteni szándékozott magánhangzó-minőség esetében, úgy, hogy ez az arány nem tekinthető elhanyagolhatónak, mert valamivel a véletlen szintje fölött van (a véletlen szintje ugyanis a nyolc megadott lehetséges válasz esetén 12,5%). Míg az /a/ válaszokat a korábbi tanulmányokban megszokott módon a szerzők a nyiltabbként ejtésből és a magasabb F_1 értékből következő nyiltabbként azonosítással, azaz az akusztikum és a percepció egyszerű lineáris összefüggéseiből magyarázzák, ez utóbbi, a nagyobb számú /u/ választ érintő eredmény értelmezésére nem tesznek kísérletet.

Mivel HOLLIEN és munkatársai (2000) vizsgálatában az akusztikai elemzésekben a kontrollfeltételt, azaz az összehasonlítás alapját nem a beszéd, hanem a „mély” kondíció adta, nem tudjuk meg, hogy ezen az alapfrekvencián voltak-e eltérések a beszélők beszélt és énekelt magánhangzói között. Annyi azonban bizonyos, hogy az alacsonyabb alapfrekvenciákon mért F_1 és F_2 értékekhez képest a „magas” kondícióban jóval magasabb F_1 és F_2 értékeket mértek a szerzők (természetesen minden magánhangzó esetében az f_0 és a $2f_0$ frekvenciaértékén, hiszen a hagyományos spektrális elemzésekkel mindösszesen ezen erősített felharmonikusok frekvenciaértéke állapítható meg, nem pedig a formánsfrekvenciák valós középértéke).

Az eddig ismertetett munkákhoz hasonlóan ez a tanulmány is hordoz némi ellentmondást az eredmények értelmezésében. Bár HOLLIEN és munkatársai (2000) látszólag úgy vélik, hogy a magánhangzóknak az ejtési szándéktól eltérő minőségű magánhangzóként történt azonosítása az akusztikumnak a magas alapfrekvencia és az F_1 hangolása²⁰ miatt megváltozott jellegeből ered, mégis a szerzők a tanulmány több pontján utalnak arra, hogy a tapasztalt alacsony azonosítási tendenciákon (a koartikulációs hatások miatt) minden bizonnyal „javítana” a mássalhangzó-környezet jelenléte. Ennélfogva nem teljesen világos, hogy a szerzők feltételezése szerint a mássalhangzó-környezet a magánhangzók ejtését vagy inkább csak azonosítását befolyásoló tényezőnek tekinthető-e.

Az énekelt magánhangzók azonosításának vizsgálatával kapcsolatosan összegzőképpen a következőket mondhatjuk el. A vizsgálatok eredményeiben közös, hogy az ejteni szándékozott magánhangzók azonosításának az alapfrekvencia emelésével csökkenő tendenciáját

²⁰ A szerzők valójában nem említik az $F_1 : f_0$ hangolást (feltehetően azért, mert a „formáns” fogalmát az akusztikai kimenetben mérhető energiamaximumra értik), így mindösszesen csak azzal kapcsolatos spekulációikat fogalmazták meg, hogy a megemelt f_0 miatt az F_1 törvényszerűen magasabb lesz, mégpedig, szintén törvényszerűen, az f_0 értékével egyező.

mutatták ki. Azon vizsgálatok, melyek a magánhangzót körülvevő mássalhangzó-környezet pozitív hatását kívánták bebizonyítani, azaz azt feltételezték, hogy az alapfrekvencia emelése következtében előálló artikulációs/akusztikai változások a formánsátmenetekben kódolt információ miatt percepciósan(?) kiküszöbölhetőek, várakozásainak megfelelő eredményeket találtak. Magyarázataikat alaposabban szemügyre véve azonban azt találtuk, hogy azok ellentmondások, és több ponton vegyítik a magánhangzók produkciós és percepciósi vetületeivel kapcsolatos érveket és feltételezéseket. E tanulmányok szerzői ugyanis nem foglalnak egyértelműen állást abban a kérdésben, hogy az F_1 -nek a megemelt alapfrekvenciára hangolását elfogadják-e a magas alapfrekvencián énekelt hangzók esetében, és ha igen, miként értelmezendő a mássalhangzók általuk javasolt „kompenzációs” hatása az ily módon (az alacsony alapfrekvenciákhoz képest) megváltozott akusztikai szerkezetű magánhangzók észlelésében.

Bár az idézett tanulmányokban a legtöbb esetben a percepciósi adatokhoz nem kapcsolódtak akusztikai elemzések, azokban az esetekben, ahol a percepciósi adatokat bármilyen jellegű akusztikai adatokkal is összevetették a szerzők, az F_1 -nek az f_0 -ra hangolásának jeleit (vagy az f_0 és F_1 „összeolvadását”) vélték felfedezni magasabb alapfrekvenciákon (vö. HOLLIEN et al. 2000), valamint a veláris és palatális magánhangzócsoportokra jellemző, eltérő akusztikai és észlelési mintázatokat (vö. BENOLKEN–SWANSON 1990).

Következtetéseik megfogalmazásában az énekelt magánhangzók percepcióját vizsgáló szerzők gyakorlatilag minden esetben valamely az artikuláció és a percepció, illetve az akusztikum és a percepció között fennálló egyszerű, direkt és lineáris összefüggés feltételezéséből indultak ki. Ezek értelmében az $F_1 : f_0$ hangolás következtében nagyobb állkapocsnnyitással, illetve magasabb F_1 -gyel ejtett magánhangzók észlelésében is az „eredeti” magánhangzó-minőségeknél nyíltabb magánhangzók felismerését tartották feltételezhetőnek. Az ezekkel az elvárásokkal egybevágó eredményeiket tehát indokoltnak (és további magyarázatot nem igénylőnek) találták, míg az ezeknek ellentmondó eredményeiket nem taglalták részletesen, amivel együtt azokra magyarázatot sem kerestek. Mindebből tehát úgy tűnik, hogy a szakirodalom javaslata szerint a magas alapfrekvenciákon énekelt magánhangzók azonosítása, az azonosításra jellemző tendenciák triviálisak. Mégis okunk van feltételezni, hogy ezek a javaslatok túlságosan leegyszerűsítőek, és az esetek egy részében legalábbis biztosan nem elegendőek a percepciósi működések értelmezéséhez.

Egyrészt, ahogyan azt az iménti ismertetések során is láttuk, az egyes vizsgálatokban találni az ezzel az egyszerű lineáris összefüggéssel nem jól magyarázható tendenciákat, valamint ellentmondásokat is, vö. pl. az /i/ és /e/ az ejtési szándéknak megfelelő azonosítása magas alapfrekvenciákon SCOTTO DI CARLO és GERMAIN (1985) vizsgálatában vagy a nagyobb arányban megjelenő /u/ válaszok HOLLIEN és munkatársainak (2000) eredményei között. Másrészt pedig nem szabad megfeledkeznünk arról sem, hogy a magánhangzók és magánhangzószzerű hangok feldolgozásának egyes kísérletes eredményei szerint a legtöbb percepciósi helyzetben nem elegendő egyetlen formáns frekvenciaértékéből kiindulnunk a beszédhangok észleleti vetületének megértéséhez. Ezen vizsgálatok értelmében ugyanis például a magánhangzók

nyíltságának észlelete (mely az énekelt magánhangzók esetében az F_1 frekvenciaértékének intenciózus módosításai miatt a legégetőbb kérdésnek látszik) nem pusztán az első formáns frekvenciaértékének, hanem inkább a spektrum prominenciái közti relációknak, elsősorban az f_0 és az F_1 viszonyának a függvénye (vö. a jelen értekezés 1.1.2.2. és 7. fejezetével).

Mindent egybevetve úgy látjuk, hogy az énekelt magánhangzók percepciója az artikulációs és akusztikai megvalósulás vizsgálatához képest erősen alulreprezentált, méltatlanul mellőzött kérdés a szakirodalomban. Ennek a hiányosságnak az egyik aspektusa, hogy az énekelt magánhangzók észleléséről eleve csak viszonylag kisszámú kísérleti eredményt találunk az irodalomban. Egy másik aspektus, hogy a vizsgálatok java az artikuláció és észlelés vagy az akusztikum és észlelés között egyszerű, direkt és lineáris összefüggéseket feltételezett, ily módon tehát ezek a vizsgálatok feltehetően túlzó módon leegyszerűsítették az énekelt magánhangzók azonosításának kérdését – miközben láttuk azt is, hogy ezek a leegyszerűsítő összefüggések több ponton nem magyarázzák a talált tendenciákat. Egy harmadik aspektus, hogy a kísérletek legtöbbször a vizsgált kérdés megválaszolásában módszertanilag több, az eredmények értelmezését is megnehezítő problémát tartalmazott. Egy negyedik aspektus, hogy a vizsgálatok nem igyekeztek az egyes nyelvekre jellemző magánhangzókészlettel szorosabb összefüggésbe hozni az énekelt magánhangzókra jellemző percepciósi tendenciákat, így valójában igen keveset tudtunk meg a percepciósi működésekről a magas alapfrekvenciájú énekelt magánhangzók azonosításában. A nem az ejtési szándék szerinti észlelési tendenciákat ugyanis a legtöbb vizsgálatban nem (vagy nem megbízhatóan) elemezték a szerzők. Egy ötödik aspektus, hogy bár a magánhangzók hangkörnyezetben ejtett megvalósításainak azonosítását és ilyen értelemben a hangkörnyezet hatását a szakirodalom megválaszolt kérdésnek látszik kezelni, úgy véljük, a problémakörben adott válaszok korántsem megnyugtatók, és további kérdéseket, alaposabb vizsgálatot érdemelnek. Végezetül pedig a hatodik szempont, hogy amint láttuk, igen ritka az olyan vizsgálat, melyben az akusztikai és percepciósi eredmények összeolvasásán keresztül képet kaphatnánk arról, hogy pontosan milyen hatások és hogyan is befolyásolják a magas alapfrekvenciákon énekelt magánhangzók azonosítását. Noha tudvalévő, hogy a magas alapfrekvenciájú magánhangzók spektrális alul-mintavételezettségének köszönhetően mind a rezonanciaadatok, mind pedig az akusztikum percepciósi vetülete a beszéd és beszédészlelés vizsgálatában megszokottnál bonyolultabb kérdés, mégis úgy véljük, hasznos volna egy szisztematikus összehasonlító vizsgálat, mely legalábbis további feltételezések megfogalmazását tenné lehetővé.

A fentiek alapján úgy látjuk, hogy máig nem tartható megválaszoltnak az a kérdés, hogy pontosan hogyan is működik az emberi beszédpercepció egy olyan extrém percepciósi helyzetben, mint a magas alapfrekvenciájú énekelt magánhangzók azonosítása. Kísérleteinket így dominánsan ennek a kérdésnek szenteltük a magyar magánhangzók esetében a magánhangzók akusztikai szerkezetének vizsgálatával kiegészítve.

1.3. A jelen kutatás célja, kérdések, hipotézisek

A jelen kötetben bemutatásra kerülő vizsgálatok célja a magas alaphfrekvenciájú éneklésben megvalósuló magyar köznyelvi magánhangzók akusztikai és percepciós jellemzőinek leírása. Kísérleteink egy részében az énekelt magánhangzók produkcióját érintő kérdésekre kerestünk választ, további kérdéseinket pedig a magánhangzók percepciójával kapcsolatban fogalmaztuk meg. Ezeket a kérdéseket az előzőekben bemutatott kísérletek nyomán felmerült problémákra és felvetésekre alapoztuk, és öt kérdéscsoportba rendeztük.

A szopránénekülésben megjelenő magánhangzók akusztikumát érintő kérdések:

1. Hogyan írhatók le a magyar magánhangzók érintő produkciós változások az alaphfrekvencia függvényében, különösen akkor, ha az f_0 magasabb, mint az adott magánhangzónak a beszédben megjelenő első formánsa? Mi jellemzi a spektrális maximumokat, és mi jellemzi a toldalékcsoében az artikuláció során megjelenő (F_1 és F_2) üregi rezonanciákat?

Mivel a kérdések megválaszolására a hagyományos, Fourier-transzformáción alapuló spektrális elemzésen túl módszertani újítások alkalmazása is szükséges, a kérdéshez kapcsolódó további megoldandó problémát jelent egy megfelelő módszertani eljárás kidolgozása és tesztelése is.

A szopránénekülésben megjelenő magánhangzók percepcióját érintő kérdések:

2. Mi jellemző az énekelt magyar magánhangzók észlelési tendenciáira? Hogyan változik az ejtési szándék szerinti azonosítás az alaphfrekvencia emelésével? Igazolható-e, hogy a magas alaphfrekvenciákon csak a nyíltabb ejtésű magánhangzókra jellemző az ejtési szándéknak megfelelő azonosítás? Mi jellemzi az ejtési szándéktól eltérő azonosításokat? Igazolható-e a produkció és a percepció közti lineáris összefüggés, azaz valóban igaz-e az, hogy az alaphfrekvencia emelésével egyre nyíltabb magánhangzóminőségek jelennek meg az észleletben?

3. Miként hat a mássalhangzó-környezet az énekelt magánhangzók azonosítására? Milyen hatása van a koartikulációs formánsátmenetekben rejlő akusztikai információnak a magánhangzók észlelésére? Miként hat (és a formánsátmeneteknél tapasztalt hatástól elkülöníthető-e) a mássalhangzó-környezet akkor, ha a mássalhangzók és a magánhangzó kapcsolatából előálló szekvenciák értelmes szavakat alkotnak (tehát az észlelést a felsőbb nyelvi szintekhez kapcsolódó feldolgozási folyamatok aktivációja is segíti)?

4. Miként hat a zöngéindítás az énekelt magánhangzók elkülönítésére/azonosítására? Alkalmazható-e az énekelt magánhangzók esetében az a hangszerek hangszínének megkülönböztetésében talált eredmény, mely szerint a hangindítás mint akusztikai kulcs jelenléte növeli a hangszín felismerésében nyújtott percepciós teljesítményt?

5. Különbözik-e bármilyen módon a naív, énektechnikailag képzetlen hallgatók és az énekesek percepciója – sikeresebbek-e az énekesek az egyes énekelt hangzók az ejtési szándék szerinti azonosításában? (A kérdés vizsgálata azt célozza, hogy a lehetőségek szerint

legalaposabban körüljárjuk a produkciós változások percepciós kompenzációjának lehetséges módzatait.)

A vizsgálandó kérdésekkel kapcsolatban három hipotézist fogalmaztunk meg:

H1: Az énekesek a magas alapfrekvenciákon énekelt magyar magánhangzók első formánsát a megemelt f_0 értékére (vagy valamivel a fölé) hangolják akkor, ha egyébként az f_0 értéke magasabb lenne, mint a magánhangzóra a beszédben jellemző első formáns értéke. Ennek értelmében az $F_1 : f_0$ hangolás megjelenése az éneklésben a magánhangzók nyíltsági fokának (illetve beszédben jellemző F_1 -értékének) függvénye.

H2: Mivel magas alapfrekvenciákon az énekelt magánhangzóknak a frekvenciaszerkezete és a spektrális felbontása is megváltozik, a magas alapfrekvencián énekelt magyar magánhangzóknak az ejtési szándék szerinti azonosítása az alapfrekvencia emelésével törvényszerűen csökken. Az f_0 és F_1 csökkenő távolsága miatt a magas alapfrekvenciákon ejtett magánhangzók akusztikai megjelenése az ejtési szándéktól eltérő azonosításokban inkább a zárt hangzók felismerését facilitálja.

H3: Mivel magas alapfrekvenciákon az énekelt magánhangzók frekvenciaszerkezete megváltozik, továbbá a magánhangzók spektrális felbontása (a környező mássalhangzókkal alkotott szonoráns hangátmenetek spektrális felbontásával együtt) csökken, a mássalhangzó-környezet a formánsátmenetekben rejlő akusztikai kulcsok révén nem fejt ki pozitív hatást a magánhangzóknak az ejtési szándék szerinti azonosítására. Ugyanez érvényes az izolált ejtésű magánhangzók zöngéindítására is. Bár úgy tűnik, hogy a mássalhangzó-környezet segítheti az ejtési szándék szerinti azonosítást akkor, ha az azonosítandó magánhangzót tartalmazó hordozó hangsorok olyan értelmes szavak, melyek száma jelentősen kisebb, mint a válaszként szóba jöhető magánhangzók száma, nem találunk ilyen hatást, ha a szóba jöhető értelmes szavak száma a magánhangzók számával együtt nagy (még ha nem is feltétlenül egyenlő).

Az itt bemutatásra kerülő vizsgálatok elsődleges célja az, hogy a szopránéneklésben megjelenő magyar köznyelvi magánhangzók akusztikus és percepciós jellemzőiről segítsen képet alkotni, megérteni a magas alapfrekvenciájú éneklésben a magánhangzók ejtését érintő szükségszerű változásokat, és megismerni a magas alapfrekvencián ejtett beszédhangok észlelésének az ejtési változásokból fakadó természetszerű nehézségeit. Emellett azonban – reményeink szerint – ennek a különleges produkciós módnak, a magas alapfrekvenciájú magánhangzóejtésnek és az ehhez a sajátos produkciós módhoz kapcsolódó észlelési folyamatoknak a vizsgálatával egyúttal általánosabb érvényű következtetéseket is megfogalmazhatunk a beszédprodukcióra (azaz a beszéd artikulációs és akusztikai szerkezetére), a beszédészlelésre, sőt közvetve a beszédmegértésre is vonatkoztatva.

2. A magyar magánhangzók ejtése és észlelése a magas alapfrekvenciájú szopránéneklésben

Az első vizsgálatban a magyar köznyelvi magánhangzók akusztikai és percepciós vetületét vizsgáltuk meg a szopránéneklésben az alapfrekvencia függvényében. Első célunk az volt, hogy feltárjuk a magas alapfrekvencián énekelt magánhangzóknek a hagyományos (Fourier-transzformáción alapuló) akusztikai elemzés segítségével kinyerhető jellemzőit, illetve hogy megállapítsuk az akusztikus jellemzőknek az alapfrekvencia emelésével megfigyelhető változásait. A második célunk az ejteni szándékozott magánhangzók észlelésében az alapfrekvencia emelésével megjelenő tendenciák vizsgálata volt.

Amint azt már a *Bevezetés*ben részletesen bemutatottuk, a formánsok hagyományos (a beszédjel Fourier-transzformáción alapuló elemzéséből kiinduló) detekciója a magas alapfrekvenciák esetében problematikus: a magas alapfrekvencián ejtett magánhangzók spektrális alul-mintavételezettsége miatt ugyanis az akusztikai kimenetben nem az ejtésekor jellemző üregi sajátrezonanciák középértéke határozható meg, hanem inkább csak az egyes rezonanciák sávszélességébe tartozó felharmonikusok frekvenciája. Szükségesnek tartjuk azonban annak alapos felmérését, hogy milyen lehetőségei vannak a hagyományos elemzésnek, valamint szükséges, hogy megállapítsuk az akusztikai kimenet maximumhelyeit az észlelés szempontjából is – az emberi beszédpercepció számára ugyanis csak ezek az információk állnak rendelkezésre az egyes magánhangzó-minőségek detektálásához.

2.1. Anyag, módszer és kísérleti személyek

A vizsgálatunk anyagát a hosszú vagy fonémakategória-váltás nélkül hosszan ejthető magyar /o a: e: i: o: ø: u: y:/ köznyelvi magánhangzók képezték. Mivel a rövid magánhangzók ejtése az éneklésben problematikus lehet – hiszen felmerülhet, hogy az énekes a hosszú magánhangzókkal való kontrasztjuk miatt a rövideket igyekszik túlságosan erősen redukált időtartamban megvalósítani, így esetleg nem a hangképzési kívánalmaknak megfelelő hangszínezettel ejteni azokat –, ezért vizsgálatukat itt is, és másutt is kihagytuk a kísérleteinkből. Ezzel együtt mégis biztosított, hogy a jelen kísérletben az összes magyar köznyelvi magánhangzó-minőség vizsgálatára módunk volt, ugyanis az ebből a felsorolásból kimaradt nyelvileg rövid magánhang-

zókot korábbi eredmények és a szakirodalomban hagyományos vélekedés alapján hosszú páruktól csak időtartamukban eltérő hangzómínőségeknek tekinthetjük (vö. 1.1.1.4. *alfejezet*).

A mássalhangzó-környezet korábbi vizsgálatok eredményei alapján feltételezhető módon megkönnyíti (legalábbis a beszélve ejtett) magánhangzók felismerését (vö. 1.1.2.2. *alfejezet*). Az egyes énekesek beszámolóiból és énektechnikai szakkönyvek alapján pedig úgy tudjuk, hogy a magánhangzót övező nazális mássalhangzók a megfelelő rezonanciaérzet biztosítása miatt megkönnyítik az énekes számára a magas alaphangfrekvenciájú hangadást (vö. Kerényi 1959; Váradi 2010). Éppen ezért a jelen kísérletben a magánhangzókat nazális hangkörnyezetben, a *mVn* hangkapcsolatban ejtve (értelmes vagy értelmetlen), egy szótagú szavakban elemeztük (*man, mán, men, mén, mín, món, mön, mún, műn*).

Az akusztikailag elemzett és az észlelési tesztben is felhasznált teszthangsorokat egy professzionális szinten éneklő szoprán énekesnő (életkora 50 év) énekprodukciójában rögzítettük (csendesített szobában, Zoom H4 típusú irányított mikrofonnal, 44,1 kHz-en, 16 biten). Az énekest arra kértük, hogy a hangsorokat számára kényelmes hangerősséggel ejtse ki kitartva (azaz változatlan alaphangfrekvenciával), egyenként közel 1,5-2 s időtartamban beszélve (átlagosan 200 Hz alaphangfrekvencián), valamint énekelve az 500, 550 és 650 Hz alaphangfrekvenciákon. Az alaphangfrekvenciákat úgy választottuk ki, hogy azokon a kísérletben részt vevő énekesnő kényelmesen, megerőltetés nélkül tudja végigénekelni a hangsorokat, és így a legmagasabb, a kísérletben elért hangmagasság 650 Hz alaphangfrekvencia volt (miközben a jelen vizsgálatban alaphangfrekvenciaként választott hangmagasságok a lineáris fizikai, azaz Hertz-értékeik szerint ekvidisztáns lépésekben növekedtek). Az intonáció pontosságának biztosításához a felvételkészítéskor minden felveendő blokk előtt referenciahangot juttattunk az énekesnő fülébe fülhallgatón keresztül.

A percepció tesztben a szoprán énekessel rögzített hangsorok magánhangzóinak azonosítását teszteltük tíz nem képzett (tehát nem a beszéd- vagy énekhanggal foglalkozó, és sem beszédtechnikailag, sem énektechnikailag nem gyakorlott) adatközlő részvételével (4 férfi és 6 nő, 21 és 25 év közöttiek, átlagéletkoruk 22 év). A tesztelők feladata az volt, hogy azonosítsák az elhangzó hangsorokban felismert magánhangzót. Ehhez egy válaszlapot biztosítottunk, melyen feltüntettük az elhangzó hangsorok mássalhangzóvázát (*m_n*), így a válaszadóknak csak a köztes magánhangzó minőségét kellett megjelölniük helyesírásban. A kísérlet előtt minden adatközlőnek felhívtuk a figyelmét arra, hogy a kísérletben értelmes és értelmetlen szavak egyaránt előfordulhatnak, így ne próbáljanak értelmes szavakat felismerni az elhangzó szótagokban, csak a magánhangzó minőségére figyeljenek. A percepció tesztben a (9 magánhangzó \times 4 alaphangfrekvencia =) 36 teszthangsort mind magánhangzó-minőség, mind pedig alaphangfrekvencia szerint randomizált sorrendben prezentáltuk a résztvevőknek disztraktor hangsorokkal vegyesen (a disztraktorok különböző mássalhangzó-környezetben beszélve vagy énekelve ejtett magánhangzók voltak), fülhallgatón keresztül.

Az akusztikai elemzés során megállapítottuk a beszélve és énekelve ejtett magánhangzók első két formánsának feltételezhető értékét, azaz detektáltuk az akusztikai kimenetben megfigyelhető maximumok frekvenciahelyeit a Praat programban (BOERSMA–WEENINK 2009).

Az adatokon statisztikai próbákat végeztünk az SPSS 13.0 program segítségével: Shapiro–Wilk-próbával ellenőriztük a normális eloszlás kritériumát, parametrikus adatok esetén egyváltozós varianciaanalízist alkalmaztunk (ANOVA), nemparametrikus adatok esetén pedig χ^2 -próbát vagy Mann–Whitney U -próbát.

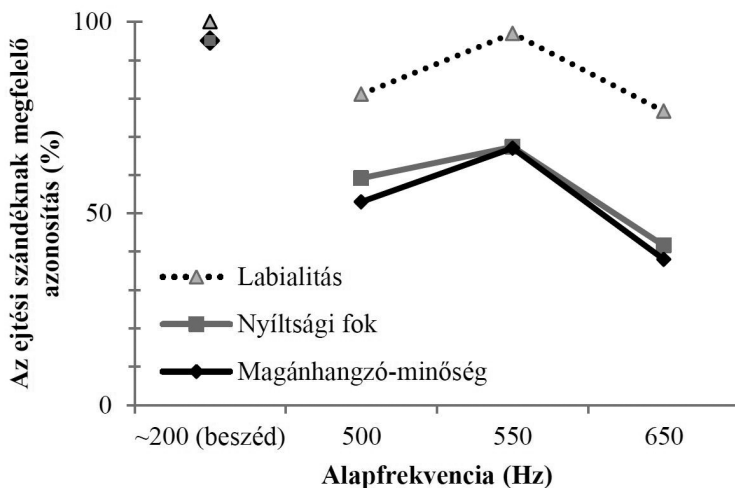
2.2. Eredmények

2.2.1. Észlelési vizsgálat

A percepciós teszt átlageredményeit a 10. ábrán foglaltuk össze. Az ábra az egyes alapfrekvenciákon ejtett magánhangzók, valamint a magánhangzók képzési jegyeinek összesített azonosítási százalékait mutatja: egy adatpont az egyes alapfrekvenciákon ejtett összes magánhangzóra adott összes válasza utal, azaz az egyes pontok a kilenc magánhangzóra a tíz adatközlő által szolgáltatott válaszokat összesítik. Az első, elkülönítve megjelenített adatpont mutatja a beszéd módra kapott azonosítási adatokat. Az elkülönítés oka, hogy ezek az adatpontok nem csak alapfrekvenciájában, hanem produkciós módjában is eltérő stimulusokra kapott eredményeket mutatnak. Az ábra tanúsága szerint az egyes magánhangzók az ejtési szándék szerinti azonosítása – az elvárásoknak megfelelően – csökkent az alapfrekvencia emelésével, tehát minél magasabb alapfrekvencián énekelte az énekes az egyes magánhangzókat, annál kevésbé ismerték fel azokat a hallgatók az ejtési szándék (azaz az énekesnek a produkciós felvételkor mutatott stimuluslista) szerint.

Az ábráról leolvasható, hogy míg beszédben a magánhangzók az ejtési szándéknak megfelelő azonosítási aránya közel 100%, addig a kísérletben a legmagasabb (650 Hz) alapfrekvencián ejtett magánhangzókat már csak 38%-ban azonosították az adatközlők az ejtési szándék szerint. Az ábra azonban egyúttal azt is mutatja, hogy az azonosítási arányok csökkenése nem teljesen fokozatos: míg 500 Hz alapfrekvencián a beszédhez képest csökkent az ejtési szándék szerinti azonosítás aránya, 550 Hz alapfrekvencián ismét megnövekedett, mígnem 650 Hz alapfrekvencián újból csökkenést mutatott.

Az eredmények szerint a labialitás jegy volt a legellenállóbb az alapfrekvencia emelésének. Ennek értelmezéséhez természetesen figyelembe kell vennünk, hogy a kilenc vizsgált magyar magánhangzó közül négy ajakréses, míg öt ajakkerekítéses, a vizsgált anyag az ajakműködés szerint tehát kiegyenlítettnek tekinthető. A labialitás és a magánhangzók nyíltsági fokának azonosítását mutató adatok összevetéséből kiderül, hogy a nyíltság a labialitásnál szignifikánsan kisebb mértékben rezisztens a hangmagasság emelésére [ANOVA, $F(2) = 8,34$; $p = 0,02$], ami azt jelenti, hogy az adatközlők az egyes énekel magánhangzókat kisebb mértékben tudták az ejtési szándéknak megfelelő nyíltsági fokú magánhangzóként azonosítani, mint az ejtési szándéknak megfelelő labialitás szerint. (A beszédben mindkét jegy azonosítása közel 100% volt).



10. ábra: Az egyes magánhangzó-minőségeknek az ejtési szándék szerinti azonosítása, valamint az ejteni szándékozott nyíltsági fok és ajakműködés az ejtési szándék szerinti azonosítása az alapfrekvencia függvényében

Az egyes magánhangzók az ejtési szándék szerinti azonosításának százalékos arányát az alapfrekvencia függvényében a 2. táblázat foglalja össze. A korábbi, az észlelést az artikulációval és percepcióval egyenes összefüggésben feltételező (az *Általános bevezetés*ben bemutatott) szakirodalom szerint az egyre magasabb alapfrekvenciák a nagyobb nyíltsági fokú magánhangzók azonosításának kedveznek – fontos felidézni azonban azt is, hogy ennek az elvárásnak egyáltalán nem minden empirikus adat felelt meg. A korábbi vizsgálatokban a lineáris összefüggésekből kiindulva nem várt eredményekkel kapcsolatban például azt jegyezték fel, hogy a magasabb alapfrekvenciákon nagy számban jelentek meg az /i:/ válaszok, az ejtési szándékknak megfelelő vagy attól eltérő válaszként egyaránt (vö. SCOTTO DI CARLO–GERMAIN 1985). A jelen vizsgálat eredményei részben szintén ezeket a korábban nem vártak tekintett tendenciákat mutatják. A 2. táblázat szerint az itt vizsgált legmagasabb alapfrekvencián nem a legnyíltsági fokú /b/ és /a:/ ejtési szándék szerinti azonosítása volt a legmagasabb, hanem az első nyíltsági fokú /i:/ és /y:/, valamint a harmadik nyíltsági fokú /e/ hangoké. A negyedik nyíltsági fokú /b/ esetében az ejtési szándék szerinti azonosítás fokozatos csökkenést mutatott, az /a:/ esetében pedig az átlagadatokban is látható ugrással megszokott, de egyébként szintén csökkenő tendenciát.

2. A magyar magánhangzók ejtése és észlelése a magas alapfrekvenciájú szopránéneklésben

2. táblázat: Az egyes ejteni szándékozott magánhangzók az ejtési szándék szerinti azonosítása az alapfrekvencia függvényében

Alapfrekvencia (Hz)	Az egyes ejteni szándékozott magánhangzók az ejtési szándék szerinti azonosításának aránya (%)								
	/a:/	/ɒ/	/o:/	/u:/	/ø:/	/y:/	/ɛ/	/e:/	/i:/
Beszéd (~200)	100	90	100	95	100	100	100	100	71
500	22	77	62	85	4	50	73	24	90
550	100	71	55	43	50	71	95	60	58
650	26	9	30	43	0	65	67	38	67

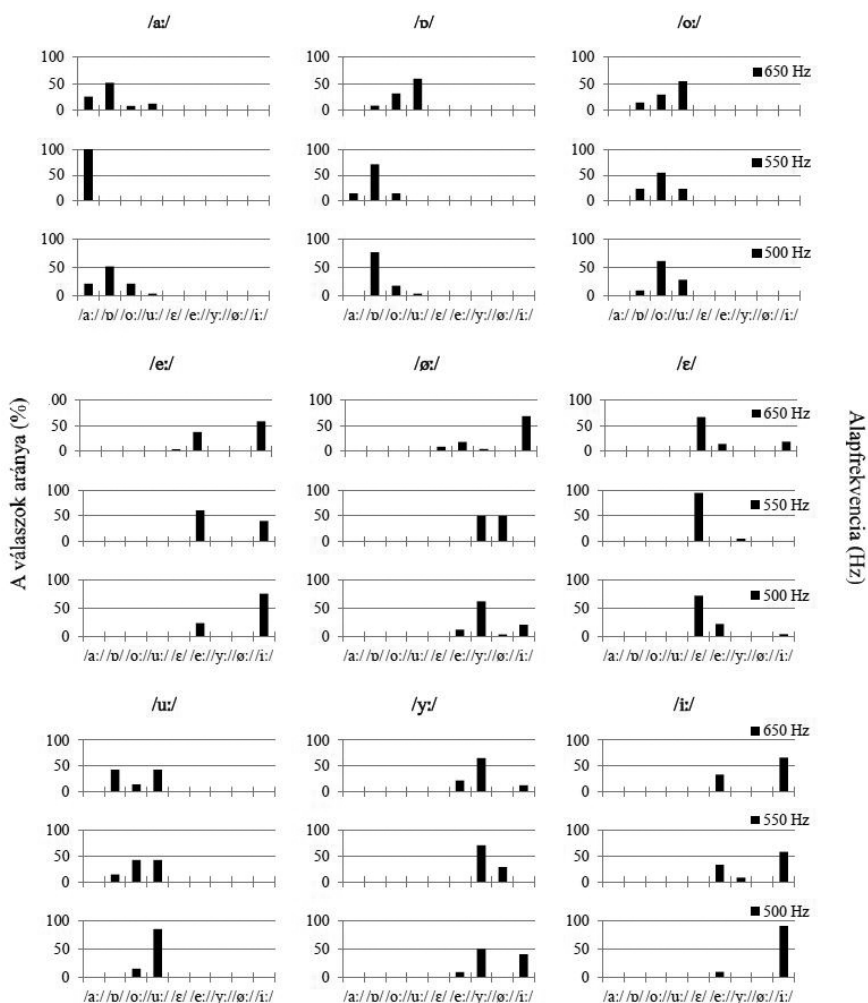
Az énekelt magánhangzók az ejtési szándéktól eltérő azonosításainak elemzéséhez előbb konfüziós mátrixokba gyűjtöttük az összes alapfrekvencián kapott összes magánhangzóválaszt (vö. 3. táblázat), majd ezeket az összesített adatokat – a könnyebb olvashatóság kedvéért – grafikusán is ábrázoltuk (11. ábra).

A 3. táblázatban példaként közölt konfüziós mátrix a következőképpen olvasható: az ejteni szándékozott /a:/ hangzót 650 Hz alapfrekvencián az adatközlők a válaszok 26,1%-ában /a/-ként, 52,2%-ban /ɒ/-ként, 8,7%-ban /o/-ként, 13%-ban pedig /u/-ként azonosították. (A mátrix átlójáról tehát az ejtési szándék szerinti azonosítások olvashatók le, míg az azon kívül eső elemek a nem az ejtési szándék szerint azonosított magánhangzóválaszokat mutatják). A 11. ábrán közölt grafikon a számadatoknak megfelelően a válaszok százalékos megoszlását ábrázolja úgy, hogy ott az adott magánhangzóhoz tartozó oszlopokban rendre az egyes alapfrekvenciákon talált eredmények szerepelnek (lentől felfelé növekvő frekvenciaérték szerint), valamint az egyes paneleken az oszlopok magassága (y érték) az x tengelyen feltüntetett magánhangzók mint válaszok megjelenésének százalékos arányát reprezentálja.

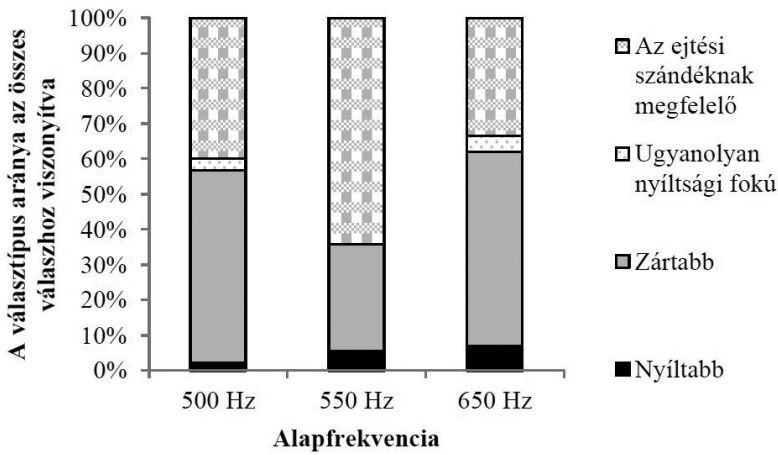
3. táblázat: A legmagasabb alapfrekvencián (650 Hz) ejtett magánhangzókra adott válaszok megoszlása a lehetséges válaszok között

		Válasz (%)								
		/a:/	/ɒ/	/o:/	/u:/	/ø:/	/y:/	/ɛ/	/e:/	/i:/
Az ejteni szándékozott magánhangzó	/a:/	26,09	52,17	11,76	13,04	0	0	0	0	0
	/ɒ/	0	9,09	31,82	59,09	0	0	0	0	0
	/o:/	0	15,00	30	55,00	0	0	0	0	0
	/u:/	0	42,86	14,29	42,86	0	0	0	0	0
	/ø:/	0	0	0	0	0	4,55	9,09	18,18	68,18
	/y:/	0	0	0	0	0	65,22	0	21,74	13,04
	/ɛ/	0	0	0	0	0	0	66,67	14,29	19,05
	/e:/	0	0	0	0	0	0	4,17	37,50	58,33
	/i:/	0	0	0	0	0	0	0	33,33	66,67

Fel kell hívnunk a figyelmet egy, mind a 3. táblázatban, mind pedig a 11. ábrán feltűnő és a korábbi szakirodalomban nem említett tendenciára. Eszerint azon hangzók esetében, melyek második vagy harmadik nyíltsági fokúak (azaz középső vagy alsó nyelvállásúak, ti. az /o: ø: e: ε/), tehát egy fokkal nyíltabb és egy fokkal zártabb szomszédjuk is van a magyar magánhangzórendszerben, gyakori, hogy az adott ejteni szándékozott magánhangzót nem az ejtési szándék szerinti nyíltsági fokú magánhangzóként azonosították a tesztelők, hanem annál zártabbként. Ennek szemléltetésére, hogy ez pontosan milyen mértékben jellemző tendencia, egy összesítő ábrát készítettünk az egyes alapfrekvenciákra a válaszok megoszlásáról a magánhangzók nyíltsági fokának tekintetében (12. ábra).



11. ábra: Az ejteni szándékozott magánhangzókra adott válaszok az alapfrekvencia függvényében



12. ábra: A második és harmadik nyíltsági fokú /e e: ø: o:/ hangzókra az egyes alapfrekvenciákon kapott válaszok megoszlása a magánhangzók nyíltsági fokának függvényében

Az énekelt magánhangzók észlelésének korábbi, az artikuláció és akusztikum, valamint a percepció lineáris összefüggéseit feltételező szakirodalma szerint az alapfrekvencia emelésével megjelenő $F_1 : f_0$ hangolás miatt a magánhangzók az ejteni szándékozott minőségükhöz képest nyíltabbnak valósulnak meg, aminek következtében a percepció rendszer is nyíltabb magánhangzó-minőségekként ismeri fel ezeket a hangzókat. A 12. ábrán közölt adatok e felvetés a lehetőségek szerint legjobban kontrollált ellenőrzését teszik lehetővé, ugyanis itt csak azon hangzókat vizsgáljuk, melyek (náluk nyíltabb és zártabb szomszédokkal is rendelkezve) mind nyíltabb, mind pedig zártabb magánhangzókként is azonosíthatóak lehetnek, annak megfelelően, hogy az alapfrekvencia emelése következtében megváltozott minőségük inkább a nyíltabb vagy inkább a zártabb szomszédjukhoz teszi őket hasonlatossá. Ez azért hangsúlyozandó, mert korábban ilyen összevetést egyetlen vizsgálatban sem végeztek, a korábbi kutatásokban az énekelt magánhangzók nyíltabbnak azonosítására vonatkozó megfigyeléseket nem ilyen jellegű összehasonlítások alapján tették.

A 12. ábrán az látható, hogy az egyes alapfrekvenciákon a kiemelt /e e: ø: o:/ hangzókra adott válaszok miként oszlanak meg akkor, ha az összevetés szempontja az ejteni szándékozott magánhangzó és a válaszként megjelölt magánhangzó nyíltsága (így az ejtési szándéktól eltérő azonosítás esetén a válasz az ejteni szándékozott hangzónál nyíltabb vagy zártabb, esetleg ugyanolyan nyelvváltsáfokú lehet). Az adatok szerint egyértelműen megállapítható, hogy az ejtési szándék szerint második és harmadik nyíltsági fokú (középső és alsó nyelvállású) magánhangzók esetében az adatközlők legnagyobbbrészt az ejteni szándékozott magánhangzónál zártabb magánhangzókat azonosítottak akkor, amikor nem az ejteni kívánt magánhangzókat vélték felismerni a kísérlet során. Ugyanakkor az is látható, hogy ez a tendencia az f_0 emelésével nem mutat egyértelmű összefüggést, és egyaránt dominánsnak látszik mindhárom énekelt alapfrekvencián.

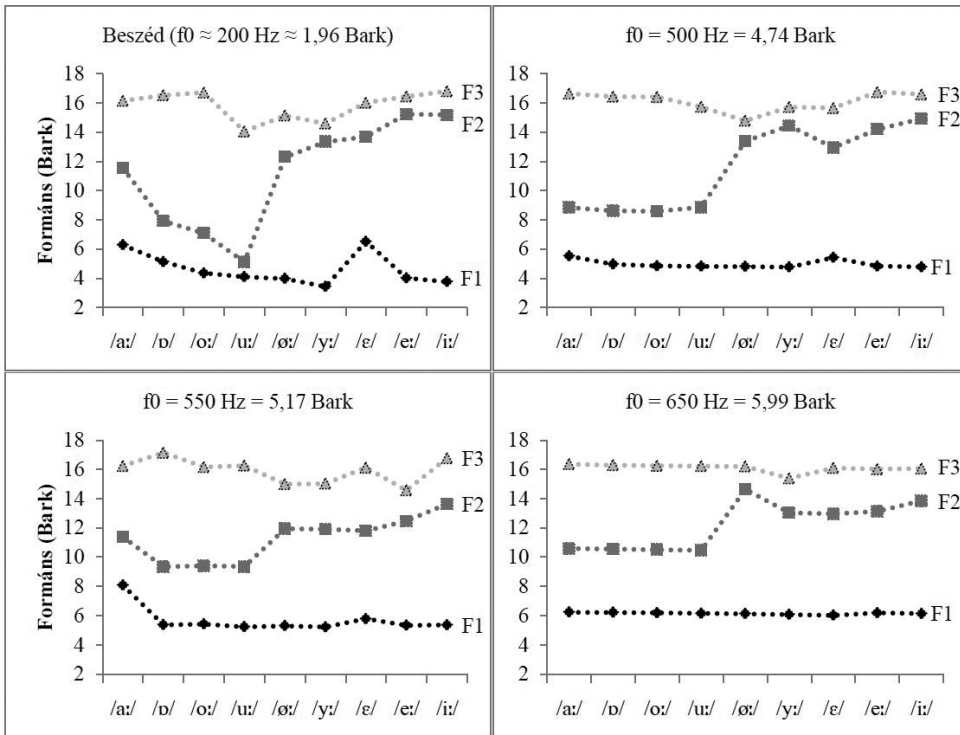
Az ejtési szándéktól eltérő azonosítások összesítésekor az is kiderült, hogy az összes stimulus-válasz eltérést figyelembe véve az /i:/ volt a leggyakoribb válaszként megjelölt hangzó, az /i:/-t az esetek 28%-ában jelölték a hallgatók az ejteni szándékozott más minőségű magánhangzók helyett. A hierarchiában az /i:/-t az /ɒ/ (16%), az /u:/ (14%), az /e:/ (14%), az /o:/ (13%) és az /y:/ (11%) válaszok követték.

2.2.2. Akusztikai elemzés

A hangnyomás-idő függvény Fourier-elemzésén alapuló akusztikai méréseink során megállapítottuk a szoprán énekesnek a beszéd módban és ének módban ejtett magánhangzóinak F_1 -, F_2 - és F_3 -értékét. Ezeket érzetileg transzformált, Bark-értékekben²¹ kifejezve a 13. ábrán közöljük.

Emlékeztetőül, amint azt már az *Általános bevezetés*ben tárgyaltuk, a formáns fogalom meghatározása alapvetően kettős: míg a kutatók egy része az akusztikai kimenetben mérhető spektrális maximumokat érti rajta, egy másik csoportjuk a toldalékcső mint rezonátorüreg sajátfrekvenciáiként, azaz rezonanciáiként definiálja (vö. 1.1.1.2. *alfejezet*). Ily módon az itt közölt adatok az első definíció szerint meghatározott formánsok méréséből származnak. A jelen vizsgálatban tehát kizárólag a mikrofon által rögzített hangnyomás-idő függvényt elemeztük, annak érdekében, hogy az akusztikai jelben mérhető, a percepciós rendszer számára elérhető spektrális energiamaximumokat megállapíthassuk. Ebből következően a 13. ábrán látható formánsadatok alacsonyabb alapfrekvencián (de elsősorban a beszédben) nagy valószínűséggel egybeesnek ugyan az artikulációs konfigurációra utaló első három rezonancia frekvenciájával, magasabb alapfrekvenciákon azonban az ábrán közölt adatok nem feltétlenül az egyes üregi rezonanciák középértékét mutatják, mint inkább az adott rezonancia sáv szélességébe tartozó legintenzívebb felharmonikus értékét. Éppen ezért ezen adatok alapján az artikulációra vonatkozó következtetések csak megszorításokkal tehetők, az adatok sokkal inkább „csak” a percepció számára elérhető akusztikus prominenciákat mutatják megbízhatóan.

²¹ A lineáris Hertz-értékek (f) érzeti transzformációját (Barkba) ZWICKER és TERHARD (1980) alapján a következő képlet szerint hajtottuk végre: $Bark = 13 \times \arctan(0,00076 \times f) + 3,5 \times \arctan((f/7500)^2)$.



13. ábra: Az énekelve és beszélve ejtett magánhangzók spektrumában mérhető első három formáns mint spektrális energiamaximum értéke (F_1 , F_2 , F_3)

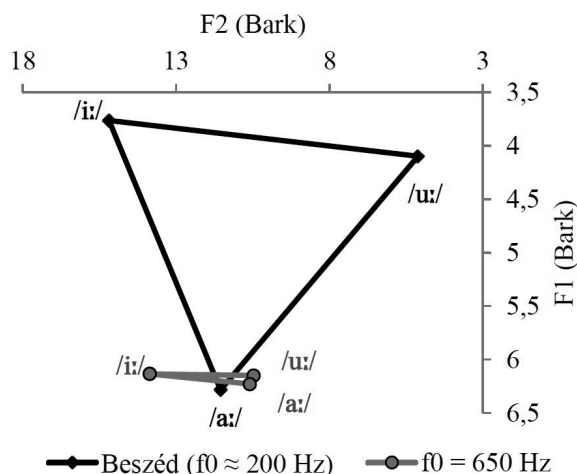
A 13. ábrán közölt akusztikai adataink szerint azokban az esetekben, amikor az alapfrekvencia emelése során az énekes hangjának alapfrekvenciája megközelítette a magánhangzó a beszédben mérhető F_1 formánsának értékét, az F_1 a megemelt (magas) f_0 értékére vagy a fölé hangolódott. Erre abból következtethetünk, hogy az f_0 és az F_1 az adott (500 Hz) alapfrekvenciától kezdve nem volt elkülöníthető a magánhangzó spektrumában, miközben a spektrum legintenzívebb akusztikai összetevője az első harmonikus, azaz az alapfrekvencia (f_0) volt (melyet feltehetőleg az F_1 üregi rezonancia erősített). Ennek következményeként az alapfrekvencia emelésének hatására az F_1 spektrális maximum értéke uniformizálódott az egyes magánhangzók között, és a magánhangzók hangszíne is egyre inkább homogenizálódott az f_0 emelésével a veláris és palatális csoportokon belül. Ez az eredmény egybecseng a korábban bemutatott percepciós adatokkal, melyek szerint a magánhangzók nyíltsági fokának az ejtési szándék szerinti azonosítása az alapfrekvencia emelésével csökken – hiszen az akusztikai eredmények is a nyíltságészlelettel szoros összefüggésben lévő F_1 formáns (vagy első spektrális maximum) eltolódását mutatják az alapfrekvencia emelésével. Az F_1 emelkedése percepciós adataink szerint azonban nem a magánhangzók észlelt nyíltsági fokának növekedését idézte elő (amint azt korábbi vizsgálatok eredményeinek értékelésekor feltételezték), hanem

sokkal inkább csak a válaszok nagyobb mértékű szóródását, tehát a magánhangzó-minőség elbizonytalanodását, amivel együtt az egyes magánhangzók az ejtési szándék szerinti nyíltsági fokánál zártabb magánhangzók azonosítása volt jellemző gyakorlatilag az összes énekelt alapprofrekvencián.

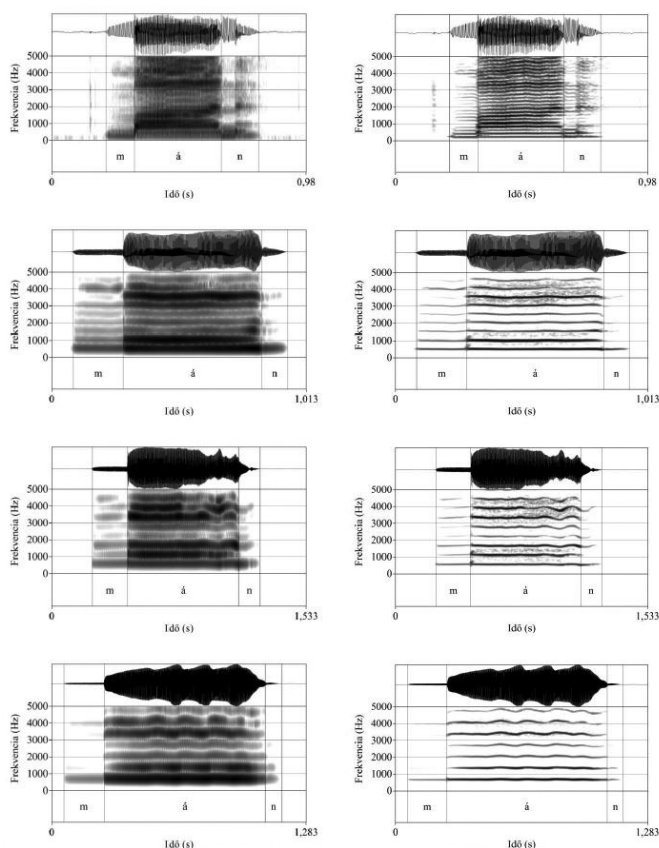
Az F_2 -értékek (azaz a második spektrális maximumhely) alakulásából úgy látszik, hogy az elől és hátul képzett hangzók közötti spektrális különbségek még a kísérletben vizsgált legmagasabb (650 Hz) alapprofrekvencián is megmaradtak – mindannak ellenére, hogy a fokozódó spektrális alul-mintavételezettség (vö. 15. ábra) miatt az F_2 mint spektrális maximum értékét a legmagasabb alapprofrekvencián már gyakorlatilag csak egyetlen egy felharmonikus határozta meg. A percepció adatok ezzel az eredménnyel is egybecsengnek: a 11. ábrán látottak szerint ugyanis gyakorlatilag egyáltalán nem fordult elő az észlelési tesztben az elől vagy hátul képzettségnek az ejtési szándéktól eltérő felismerése.

Végezetül összevetettük az artikulációs és akusztikai tekintetben egymástól legmesszebb elhelyezkedő /a: i: u:/ hangok által közrezárt (azok spektrális maximumai által meghatározható) akusztikai magánhangzóteret a beszédben és a kísérletben elért legmagasabb énekelt alapprofrekvencián (14. ábra).

A várakozásoknak megfelelően az akusztikai magánhangzóter az alapprofrekvencia emelésével jelentősen redukálódott, és eltolódott az /a:/ pozíciója felé, ugyanakkor jól láthatóan 650 Hz alapprofrekvencián a magánhangzóter az elől és hátul képzettség dimenziójában (F_2) nem csökkent le olyan mértékben, mint a nyíltságnak szokásosan megfeleltethető F_1 dimenzió mentén.



14. ábra: Az akusztikai tekintetben egymástól legmesszebb elhelyezkedő /a: i: u:/ magánhangzók által meghatározott akusztikai magánhangzóter a beszédben és a kísérletben elért legmagasabb énekelt alapprofrekvencián



15. ábra: A *mán* hangsor beszédben ejtett (első sor), valamint az 500, 600 és 650 Hz alapfrekvencián énekelt (2–4. sor) hullámformája (az egyes panelek felső részében), valamint széles sávú és keskeny sávú hangszíne (az egyes panelek alsó részében, rendre a lap bal és jobb oldalán)

A 15. ábrán a *mán* hangsor hullámformája, valamint széles sávú (bal oldal) és keskeny sávú (jobb oldal) spektrogramja látható. Az első sorban a beszédbeli megvalósítást találjuk, a másodiktól a negyedik sorig pedig (rendre) az 500, 550 és 650 Hz alapfrekvenciákon énekelt hangsorokat. A spektrogramokon (de különösen a keskeny sávú spektrogramon) jól kivehetőek a harmonikusok (az alaphang és annak felharmonikusai, melyek egymástól egyenlő távolságokra elhelyezkedő vízszintes csíkként jelennek meg), így az is megfigyelhető, hogy az alapfrekvencia emelésével miként kerülnek ezek egymástól egyre nagyobb és nagyobb távolságra. Mivel a felhangok intenzitására ebben a spektrografikus megjelenítésben a sötétedés mértéke utal, ezért a toldalékcso rezonanciái (legalábbis a beszédjelben) sötétebb színű, szélesebb sávokként jelennek meg a spektrogramon. Jól látható, hogy míg a beszédjelben a felerősített felharmonikus-nyalábok egyértelműen kirajzolják az egyes (sötétebb sávokként megjelenő) rezonanciahelyeket, azaz a formánsokat, addig a magasabb alapfrekvenciákon

a formánsok frekvenciahelyének beazonosítása már sokkal bizonytalanabb. Ezeken az alapprofrekvenciákon ugyanis az intenzitásérősödések már sokkal inkább csak az egyes felharmonikusok erősítettségére utalnak megbízhatóan, a formánsok középértékére azonban nem – „formánsértékeként” az akusztikai kimenetből tehát csak az erősített felharmonikusok frekvenciaértékei, nem pedig az üregi rezonanciák középértéke detektálható.

2.3. Következtetések

Az itt ismertetett vizsgálat az első olyan kísérletsorozat, mely mind produkciós, mind pedig percepciós szempontból igyekezett feltérképezni az éneklés mint produkciós mód hatását egy adott nyelv összes magánhangzójára.

Percepciós vizsgálatunk eredményeit összegezve megállapítható, hogy kísérletünkben nem igazolódott az a szakirodalomban gyakorta olvasható (de alapvetően csak spekulatív) feltételezés, mely szerint az alapprofrekvencia emelésével (az $F_1 : f_0$ hangolás következtében) egyre magasabb F_1 a megvalósult magánhangzók percepciós nyitódását okozná. Ezt a megállapítást arra a jól kontrollált összevetésre alapozhatjuk, amelyben a nyíltabb és zártabb szomszédokkal rendelkező második és harmadik nyíltsági fokú magánhangzókra kapott válaszokat elemeztük. Eszerint ugyanis magas alapprofrekvenciákon (már a kísérletben legalacsonyabb énekelt alapprofrekvenciaként vizsgált 500 Hz-on és a fölött is) a legjellemzőbb észlelési tendencia az egyes magánhangzók az ejtési szándéktól eltérő azonosításában a zártabb hangzóként való felismerés volt. A jelen vizsgálattal szemben a korábbi vizsgálatokban vagy nem derül ki pontosan, mire alapozták a szerzők a percepciós nyitódásra vonatkozó megállapításukat (vö. SCOTTO DI CARLO–GERMAIN 1985), vagy eleve csak olyan zárt (első nyíltsági fokú) magánhangzókat vizsgáltak, melyek – nem lévén zártabb szomszédjuk – „tévesztés” esetén eleve csak nyíltabb hangzókként voltak azonosíthatók (vö. HOLLIEN et al. 2000). Éppen ezért a korábbi eredményeket a jelen vizsgálat eredményeivel nem volna érdemes ütköztetnünk. Az ejtési szándéktól eltérő azonosításokról, valamint az alapprofrekvencia emelésével megjelenő észlelési tendenciákról legáltalánosabban annyit mondhatunk, hogy az f_0 emelésével egyre csökken az egyes magánhangzók az ejtési szándék szerinti azonosításának aránya, miközben a válaszok szóródása egyre növekszik – ugyanis a válaszként adott magánhangzók egyre nagyobb szóródást mutatnak a különböző magánhangzó-minőségek között.

A kísérletben nem igazolódott az a korábbi szakirodalomban megfogalmazott feltételezés sem (vö. pl. SCOTTO DI CARLO–GERMAIN 1985), mely szerint a magasabb alapprofrekvenciákon a nyíltabb magánhangzók az ejtési szándéknak megfelelő felismerése volna a legjobb. Bár a kísérletben elért 650 Hz alapprofrekvenciát a szoprán hangterjedelem szempontjából nem tekinthetjük kiemelkedően magasnak (hiszen a szoprán hangterjedelem felső határának általában a „magas C” vagy a c''' zenei hangot, azaz az 1047 Hz alapprofrekvenciát tekintik, lásd pl. MCKINNEY

1994), mégis ezen a – fogalmazzunk így – közepesen magas alapfrekvencián a legmagasabb az ejtési szándék szerinti azonosítást a zárt, első nyíltsági fokú /i:/ és /y:/ mutatta, valamint a harmadik nyíltsági fokú /ɛ/. Ezzel párhuzamosan az /i:/ más szempontok szerint is kiemelt szerepét láttuk: az /i:/ volt ugyanis a leggyakoribb válasz az ejtési szándéktól eltérő azonosítások során.

Az /i:/ különleges percepciós tulajdonságait már több vizsgálatban is kimutatták. Az éneklésre SCOTTO DI CARLO és GERMAIN (1985) jegyezte le először az /i:/ mint gyakori válasz megjelenését a magasabb alapfrekvenciákon, míg a beszédre MÁDY és REICHEL (2007) is kimutatta azt az összefüggést, mely szerint a hosszan ejtett /e:/ magánhangzóra gyakran a nála zártabb minőségű /i:/ válasz érkezett (míg más magánhangzók hosszú ejtésénél nem volt jellemző a hangzók zártabbként való azonosítása). A SCOTTO DI CARLO és GERMAIN (1985) által csak a korábbi artikulációs vizsgálatok alapján feltételezett, MÁDY és REICHEL (2007) által pedig akusztikai mérések alapján megállapított formáns adatok ugyanakkor egyik esetben sem magyarították ezeket a „tévesztéseket”. Szintén említhető itt GÓSYNAK (1989) a magánhangzók azonosítását vizsgáló kísérlete, melyben a szerző különböző szűrők alkalmazásával nehezítette a magánhangzók a beszédészlelés számára fontosnak vélt akusztikai kulcsainak percepciós elérését. Ebben a vizsgálatban ugyanis azt találta a szerző, hogy az /i:/ és /y:/ magánhangzók azonosítási aránya a csak igen magas frekvenciájú összetevőket tartalmazó frekvenciasáv (2700–3300 Hz) alapján is viszonylag magas volt (vö. *I.1.2.3. alfejezet*).

Az /i/ hangzót RYALLS és LIBERMAN – korábbi percepciós vizsgálatokra is hivatkozva – „szupermagánhangzónak” nevezi (1982: 1633), mégpedig azért, mert az azonosítást tekintve az összes általuk vizsgált magánhangzó közül az /i/-t találták a legkevésbé érzékenynek a különböző „nehezített” percepciós feladatokban. Magyarázatauk szerint az /i/ kitüntetett szerepét különös szalienciájának köszönheti, distinktív akusztikai jegyei ugyanis minden más magánhangzónál jobban elkülöníthetővé teszik a magánhangzórendszer (vagy talán inkább a magánhangzótér) többi tagjától. Akárhogy is (talán ezekkel a korábbi eredményekkel összefüggésben, talán nem), a jelen percepciós vizsgálatban igen erős /i/-preferenciát találtunk a válaszáadásban, melynek magyarázata pusztán az egyes énekelt magánhangzók megvalósulásának akusztikai vetülete (és az F_1 és F_2 mint spektrális maximumok abszolút értéke) alapján nem tűnik lehetségesnek.

A magas alapfrekvencián énekelt magánhangzók akusztikai vetületében a magas alapfrekvencia (és a magas alapfrekvencián ejtett magánhangzók spektrális alul-mintavételezettsége) miatt elsősorban nem a toldalékcsőre mint üregrendszerre az artikuláció során jellemző rezonanciák, hanem a megvalósított hangzókból tapasztalható energiamaximumok mérhetők, melyek megfelelően magas f_0 esetén sokszor nem egyeznek a toldalékcső rezonanciahelyeivel. A jelen fejezetben ismertetett mérések során ennek megfelelően nem az artikulációt többé-kevésbé pontosan leíró rezonanciafrekvenciákat, hanem a kisugárzott hangspektrum energiamaximumainak frekvenciahelyeit detektáltuk. Ez az akusztikai információ azért bír jelentőséggel, mert a beszédészlelés számára rendelkezésre álló prominenciákra utal, ugyanakkor ezekből az adatokból a magánhangzók létrehozásakor jellemző artikulációs működésekre csak nagyon áttételesen következtethetünk.

Az akusztikai kimenetben mért spektrális maximumok vizsgálata alapján arra következtethetünk, hogy valószínűleg fellépett az $F_1 : f_0$ hangolás azokban az esetekben, amikor a megemelt f_0 frekvenciaértéke egyébként magasabb lett volna, mint az adott hangzónak a beszédben jellemző első formánsa. Ezt főként arra a megfigyelésre alapozzuk, hogy magasabb alapfrekvenciákon minden esetben az f_0 volt a legintenzívebb frekvenciakomponens, melytől az F_1 elkülöníthetetlennek bizonyult. Az F_2 mint spektrális maximum ezzel szemben a 650 Hz alapfrekvencia eléréséig nem uniformizálódott teljesen a magánhangzók között (még annak ellenére sem, hogy a spektrális alul-mintavételezés miatt csökkent az F_2 -höz mint rezonanciához eső, a spektrális maximumot kialakító felharmonikusok száma). Így a palatális és veláris magánhangzócsoportok az F_2 mentén még az itt vizsgált legmagasabb alapfrekvencián is elkülönültek egymástól mind akusztikai, mind pedig perceptuális tekintetben (ez utóbbi következtetést az észlelési teszt eredményeire alapozzuk).

Az akusztikai magánhangzótér az F_1 és F_2 formáns (mint spektrális maximumhely) változásának következtében a legmagasabb énekelt alapfrekvencián (a beszédhez képest) jelentősen redukálódott, és a nyílt, centrális /a:/ pozíciója felé tolódott. Érdekes kérdés, hogy ennek ellenére miért nem növekedett meg drasztikusan az /a:/-jelölések száma a válaszok között. A kérdés megválaszolásához azonban feltehetőleg több kell, mint az egyes megvalósított hangzók első két spektrális prominenciájának itt mért akusztikai adatai, valamint annak a feltételezése, hogy az artikuláció és az akusztikum a percepcióval egyszerű, lineáris összefüggésben áll. Mindezek helyett valószínűleg gyümölcsözőbb volna azokat a percepciós eredményeket és modelleket segítségül hívni az itt tapasztalt észlelési folyamatok megértéséhez, melyek a zártságot az f_0 és az F_1 tonotópiás viszonyának függvényében értelmezik. TRAUNMÜLLER (1981) eredményei szerint a nyíltság érzeti dimenziójában az F_1 abszolút frekvenciaértéke helyett sokkal inkább az f_0 és F_1 távolsága a döntő (vö. a jelen értekezés *1.1.2.2. alfejezete*), mely az éneklésben az alapfrekvencia emelésével minden bizonnyal változó paraméter. Az $F_1 : f_0$ hangolás megléte vagy hiánya azonban az elmélet alkalmazása szempontjából kritikus kérdés, ahogyan az F_2 értéke is – az elmélet alkalmazásához ugyanis szükséges lenne (a jelen adatoknál megbízhatóbban) tudnunk, hogy pontosan mely formánsok, és mely felharmonikusokat erősítik. A fejezetben bemutatott adatok alapján ezért (egyelőre) kérdéses marad, miként volna adaptálható a modell az éneklésben tapasztalt észlelési tendenciákra; és csak sejthető, hogy az F_1 és az f_0 közti csökkenő tonotópiás távolság lehet a felelős a magánhangzóknak az ejtési szándéknál zártabb észleletére. A későbbiekben azonban, a *7. fejezetben* bemutatott kísérlet során – melyben az észlelési adatokat a rezonanciaadatok tükrében vizsgáljuk –, még ennél részletekbe menőbb módon is megkíséreljük az elmélet adaptációját. Addig is (a jelen vizsgálat fontos konklúziójaként) kiemelendő az a jelen vizsgálatból nyert egyértelmű empirikus bizonyíték, mely szerint a pontosan kontrollált és releváns összehasonlításokat tartalmazó percepciós vizsgálat nem utal a produkció és az észlelés közötti egyszerű, lineáris összefüggésekre, melyeket korábbi vizsgálatok során egyértelműen elfogadhatónak tartottak.

Abból, hogy az alapfrekvencia emelésével egyre csökkent a magánhangzók az ejtési szándék szerinti azonosítása, valamint abból, hogy a válaszok szóródása az egyre magasabb f_0 elérésével egyre nagyobb lett, az következik, hogy az énekelt magánhangzók akusztikai és percepciósi minősége az alapfrekvencia emelésével együtt változott. E változás természete az eredmények szerint nem kategoriális, azaz az adatközlők nem egy észleleti kategóriáról hirtelen a másikra váltva azonosították az egyes hangzókat, hanem fokozatos, aminek következtében a magánhangzók percepciósi minősége (legalábbis 650 Hz alapfrekvenciáig) az alapfrekvencia emelésével egyre bizonytalanabbnak tűnt fel. Mindezek fényében (az *1. fejezet* végén megfogalmazott hipotéziseinkre is utalva) úgy véljük, kérdéses, hogy várható-e bármilyen pozitív hatás a mássalhangzó-környezettől mint a koartikulációs formánsátmenetek révén a közbeeső magánhangzó minőségét kulcsoló percepciósi „mankóktól”. Ha ugyanis a magánhangzók minősége valóban megváltozik, a formánsátmenetek is legfeljebb csak erre a megváltozott magánhangzó-minőségre (nem pedig egy „elméleti”, de meg nem valósuló magánhangzó-minőségre) utaló kulcsokat tartalmazhatnak. A kérdésre az *5. fejezet*ben bemutatott kísérletekkel keressük a választ.

További (részben az előzőhöz kapcsolódó) kérdésként merül fel, hogy vajon hatással volt-e a jelen kísérletben használt mássalhangzó-környezet minősége, azaz a két nazális mássalhangzó az észlelési eredményekre. Az internazális magánhangzó nazalizációja ugyanis megváltoztatja a magánhangzók akusztikai szerkezetét, az akusztikai jellemzők között is leginkább az F_1 értékét (lásd KRAKOW et al. 1987 és hivatkozásai; ROSNER–PICKERING 1994; HORVÁTH 2005, 2008). Ez alapján tehát logikusan feltételezhető lenne, hogy a nazalizáció a magánhangzók észlelt nyíltságára is hatást gyakorolhat a beszédben. Ugyanakkor egyes vizsgálatok szerint e hatást, ha természetes, tehát bejósolható helyen fordul elő a beszédben (azaz akkor, ha egy magánhangzó nazális, nem pedig orális mássalhangzók mellett/között nazalizálódik), a percepciósi rendszer képes kompenzálni, így nem változik a magánhangzó zártságának észlelete (BEDDOR 1983; KRAKOW et al. 1987). Vitatott kérdés, hogy az énekelt nazális mássalhangzók a beszédben ejtettekkel egyező mértékben nazalizáltak-e, így nincs egyértelmű állásfoglalás abban a kérdésben sem, hogy az éneklésben megvalósuló nazális mássalhangzók a beszédben tapasztalhatóhoz képest milyen mértékű nazalizációt fejtenek ki a szomszédos magánhangzókra (vö. VENNARD 1964; SCOTTO DI CARLO–GERMAIN 1985; YANAGISAWA et al. 1990; YANAGISAWA et al. 1991; GRAMMING et al. 1993; STEPHEN 1996). Éppen ezért fontosnak tartjuk megvizsgálni ezt az aspektust is, és kizárni annak a lehetőségét, hogy a jelen vizsgálat eredményeit a magánhangzók nazalizációja eredményezte volna.

Az itt felvetett problémák, azaz a nazális mássalhangzó-környezet hatásának ellenőrzésére, valamint a mássalhangzó-környezetnek az izolált ejtéssel összehasonlításban tapasztalható hatásának vizsgálatára terveztük a *3. és 4. fejezet*ekben bemutatásra kerülő kísérleteket. A *6. fejezet*ben, amint azt már említettük, a formánsok mint üregi rezonanciák vizsgálatát kíséreljük meg, mely adatokat a *7. fejezet*ben egy észlelési vizsgálat eredményeivel is összevetjük.

3. A mássalhangzó-környezet hatása az énekelt magánhangzók észlelésére: a mássalhangzó képzésmódjának hatása

Az előző fejezetben a magyar köznyelvi magánhangzók akusztikai és percepciós tendenciáit vizsgáltuk az éneklésben mint különleges produkciós sajátosságokkal rendelkező beszédhang-képzési módban. Az észlelési eredmények szerint az alapprofrekvencia emelésével a magánhangzók az ejtési szándék szerinti azonosításának aránya csökkent, miközben olyan, az ejtési szándéktól eltérő azonosítási tendenciákat találtunk, melyek nem támasztják alá azokat a korábbi szakirodalomban sugallt feltételezéseket, amelyek a produkció és percepció közti egyszerű, lineáris összefüggéseken alapulnak.

Annak ellenőrzésére, hogy eredményeinket valóban elsősorban az éneklés mint sajátos artikulációs és akusztikai jellemzőkkel bíró produkciós mód hatására kaptuk, és nem más körülmények, például a választott nazális mássalhangzó-környezet miatt, a jelen vizsgálatban kibővítjük a korábbi észlelési tesztünket: a nazális hangkörnyezet mellett ez alkalommal zöngés és zöngétlen frikatívákat is tesztelünk hordozó hangsorokként.

Az orális magánhangzók nazális mássalhangzók környezetében a koartikuláció következtében nazalizálódhatnak, azaz a mássalhangzó képzésére jellemző nazális jelleg átterjedhet a nazális mássalhangzóval szomszédos magánhangzóra. A nazalizáció eredményeként az eredetileg orális minőségű magánhangzók minősége megváltozik; a szakirodalom szerint a legjelentősebb változás az F_1 formánst érinti, mégpedig a következőképpen. Korábbi akusztikai mérések szerint a nazalizáció hatására a felső nyelvállású vagy első nyíltsági fokú magánhangzók, valamint egyes második vagy harmadik nyíltsági fokú magánhangzók F_1 -értéke nő, azaz a magánhangzók akusztikailag nyílnak, míg a többi második és harmadik nyíltsági fokú magánhangzó, valamint a negyedik nyíltsági fokú magánhangzók F_1 -értéke csökken, azaz ezek a magánhangzók akusztikailag záródnak. Mindeközben pedig a nazalizált magánhangzó első spektrális prominenciája minden esetben nagyobb sávzélességű és kisebb intenzitású, mint az orális realizációé (KRAKOW et al. 1987).

Az akusztikai változások következményeként sok kutató azt feltételezte, hogy a nazalizált magánhangzók az orális magánhangzókként megvalósított megfelelőjükhöz képest kevésbé felismerhetők a beszédben. Erre a feltételezésre például WRIGHT (1975) szolgáltatott az elsők között kísérletes bizonyítékokat izolált ejtésű nazalizált magánhangzók vizsgálatával. Később azonban KRAKOW és munkatársai (1987) arra is fényt derítettek, hogy a nazalizáció csak a hangkörnyezetre vonatkozó információk hiányában hat negatívan az ejtés szerinti azonosításra, a szerzők ugyanis azt találták, hogy a nazális mássalhangzók között ejtett magánhangzók azonosítása az orális mássalhangzók között ejtett magánhangzókéhoz nagyon hasonlóan

alakult. Ebből KRAKOW és munkatársai (1987) arra következtettek, hogy bár a nazalizáció megváltoztatja a magánhangzó akusztikai szerkezetét, hatása jósolható helyen, azaz nazális mássalhangzók környezetében jelentkezve percepciósan „korrigálható”.

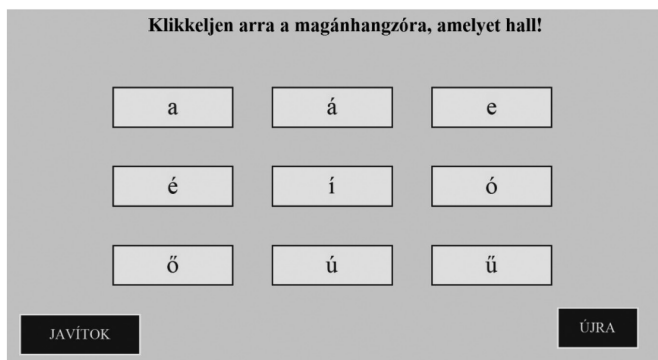
Mivel a megfelelő technikával képzett énekhang sok énekes tapasztalata szerint orrüregi rezonanciaélménnyel jár (erre utal az „előre helyezett”, „elő képzett” hang gyakran használt metaforája is), az éneklés szakirodalmában sokat kutatott, de maig vitatott kérdés, hogy vajon valóban megjelenik-e a nazalizáció az éneklésben, és ha igen, milyen mértékű lehet. Egyes kísérletek tanúsága szerint az éneklésben a nazális minőségűnek szánt mássalhangzók nazalizációjának mértéke és a szomszédos magánhangzóra kifejtett nazalizációs hatás – főként magasabb alapfrekvenciákon – jóval kisebb, mint a beszédben, vagy akár teljesen el is tűnhet (VENNARD 1964; YANAGISAWA et al. 1990; YANAGISAWA et al. 1991). Mások szerint azonban a lágy szájpad az éneklésben folyamatosan nyitva van, ráadásul úgy, hogy ebben az artikulációs stratégiában nagyfokú egyéni variabilitás is tapasztalható, miközben az orrüregi csatolás mértéke is igen változatos lehet, mind az adott magánhangzó, mind pedig a beszélő személy függvényében (GRAMMING et al. 1993). Mindemellett fontos azt is megjegyeznünk, hogy az énekelt magánhangzók nazalizációjának percepció vetületéről a jelen értekezés szerzőjének tudomása szerint még nem született vizsgálat, nem tudjuk tehát, hogy ha van is orrüregi csatolás az éneklés során, az egyáltalán érzékelhető-e bármilyen mértékben a hallgatók számára.

A fentiekből következően nem zárható ki tehát teljesen, hogy az előző kísérletünkben vizsgált magánhangzók akusztikai megvalósulását befolyásolta a szomszédos nazális mássalhangzó, így a magánhangzók percepciója is – valamilyen formában – szintén a nazalizáció hatása alatt állt. Éppen ezért a jelen fejezetben bemutatott vizsgálat célja annak a kérdésnek az ellenőrzése, hogy az előző vizsgálatban talált magánhangzó-azonosítási tendenciák mennyiben írhatók a nazális hangkörnyezet számlájára. Ehhez nazális mássalhangzó-környezetben ejtett magánhangzók azonosítását vetettük össze zöngés és zöngétlen posztalveoláris frikatívák között ejtett magánhangzók azonosításával. Mivel a szakirodalom alapján tudjuk, hogy a nazalizált magánhangzók nazális hangkörnyezetben gyakorlatilag azonos százalékban azonosíthatók, mint az obstruens környezetben ejtett megfelelőjük (vö. KRAKOW et al. 1987), a vizsgálat elején azt feltételeztük, hogy a nazális és frikatíva hangkörnyezetek hatása között nem találunk jelentős különbségeket az éneklésben sem. Ez a feltételezés tehát azonos alapszik, hogy ha igaz is, hogy az éneklésben nazális mássalhangzó szomszédságában nazalizálódik az egyébként orális magánhangzó, és ennek következményeképp annak akusztikai szerkezete is megváltozik, ez a magánhangzó hangkörnyezetben történő bemutatásakor nem változtat jelentősen a magánhangzó észlelt hangszínén.

3.1. Anyag, módszer és kísérleti személyek

Az észlelési kísérlethez az előző vizsgálatban is részt vett szoprán énekesnővel (életkora 50 év) készítettünk újabb felvételeket (csendesített szobában, Zoom H4 típusú irányított mikrofonnal, 44,1 kHz-en, 16 biten). Az énekes ismét az /*p a: ε e: i: o: ø: u: y:/* magánhangzókat ejtette ki beszéd módban (≈ 200 Hz alapfrekvencián), valamint ének módban, utóbbi esetben az 500, 550, 600, 650 Hz alapfrekvenciákon, statikus dallamívvel, azaz kitartva. Az énekelt stimulusok felvételekor ismételten a megfelelő f_0 -értéken megszólaló referenciahangokkal segítettük a pontos intonációt. Minden hangsort körülbelül 1,5-2 másodpercig hangoztatott az énekes. A magánhangzókat ez alkalommal a *mVn* valamint a *sVs* és a *zsVzs* hangkörnyezetekben vettük fel (pl. *mán, sás, zsászs* stb.). Így összesen (3 kontextus \times 9 magánhangzó \times 5 alapfrekvencia \times 2 ismétlés \Rightarrow 270 stimulus) teszteltünk disztraktor hangsorokkal kiegészítve (a disztraktorok más hangkörnyezetekben ejtett magyar magánhangzók voltak).

Az észlelési tesztben 15 naiv (tehát sem beszédtechnikailag, sem énektechnikailag nem képzett) adatközlő vett részt (10 nő és 5 férfi, 21–29 évesek, átlagos életkoruk 23 év). A percepciósi tesztet egy erre a célra a Praat programban (BOERSMA–WEENINK 2009) írt szkript vezérelte. Az adatközlők feladata az volt, hogy az egyes stimulusok elhangzása után kattintsanak a képernyőn helyesírásban bemutatott magánhangzók közül arra, amelyiket a hangsorban hallani vélték (16. ábra). A teszt előtt felhívtuk rá a kísérleti személyek figyelmét, hogy értelmes és értelmetlen egy szótagú szavak egyaránt előfordulhatnak a tesztben, ezért minden esetben a magánhangzó minőségére figyeljenek, ne az elhangzó szót vagy álszót próbálják azonosítani. A résztvevőknek lehetősége volt egyszer újrakattintani az aktuális hangsort (az *Újra* gomb megnyomásával), valamint egyszer javítani az esetleges félrekattintásból származó eredményt (a *Javítok* gomb megnyomásával). Ez után azonban mindenképpen választaniuk kellett. A tesztet a résztvevők egyesével végezték el, a stimulusokat fülhallgatón keresztül hallották.

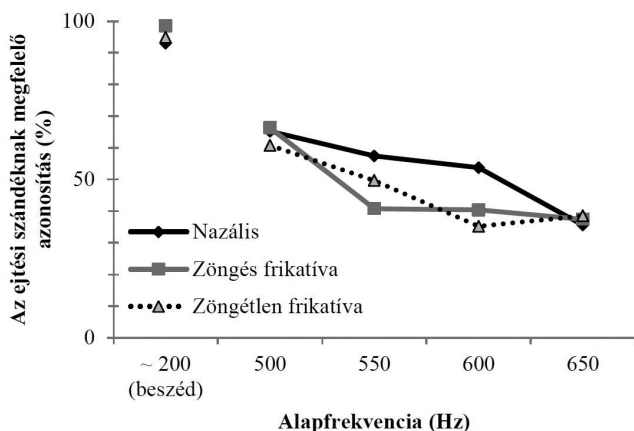


16. ábra: A percepciósi teszt felülete a Praat programban

Az észlelési eredményeken statisztikai próbákat végeztünk az SPSS 13.0 programmal: előbb Shapiro–Wilk-próbával elemeztük a normális eloszlás kritériumát, majd ennek teljesülése esetén egyváltozós varianciaanalízist (ANOVA), nem normális eloszlás esetén pedig nem-parametrikus tesztekkel (χ^2 -próba, Mann-Whitney U -próba) alkalmaztunk.

3.2. Eredmények

A beszédben és az éneklésben tapasztalt, az alapfrekvencia emelésével megjelenő átlagos azonosítási tendenciákat összegzi a 17. ábra. Az eredmények szerint a mássalhangzó-környezet minősége nincs szignifikáns hatással a magánhangzók azonosítására [ANOVA, $F(2) = 0,065, p = 0,937$]. Ugyanakkor az is látható, hogy a nazális kontextus 550 Hz és 650 Hz alapfrekvencián valamivel magasabb, az ejtési szándéknak megfelelő azonosítási arányokat eredményezett, mint a zöngés és zöngétlen frikatívák. 500 Hz és 600 Hz alapfrekvencián, valamint a beszédben gyakorlatilag nem látunk különbséget a kondíciók között.



17. ábra: Az énekelt magánhangzók azonosítása nazális, valamint zöngés és zöngétlen frikatíva környezetében az alapfrekvencia emelésével

Az ejtési szándéktól eltérő azonosítási tendenciák megfigyeléséhez a stimulus-válasz párokat a korábbiakhoz hasonlóan ismét konfúziós mátrixokba rendeztük. A legmagasabb alapfrekvencián tapasztalt azonosítási tendenciákat e mátrixokba rendezve a 4. táblázatban közöljük. A mátrixokban az 50%-os és az a feletti azonosításokat besötétítettük, hogy a kimagaslóan magas arányban az ejtési szándék szerint azonosított magánhangzókot összevethessük az egyes kontextusokban (ehhez az 50%-os határt önkényesen választottuk). Amint az látható,

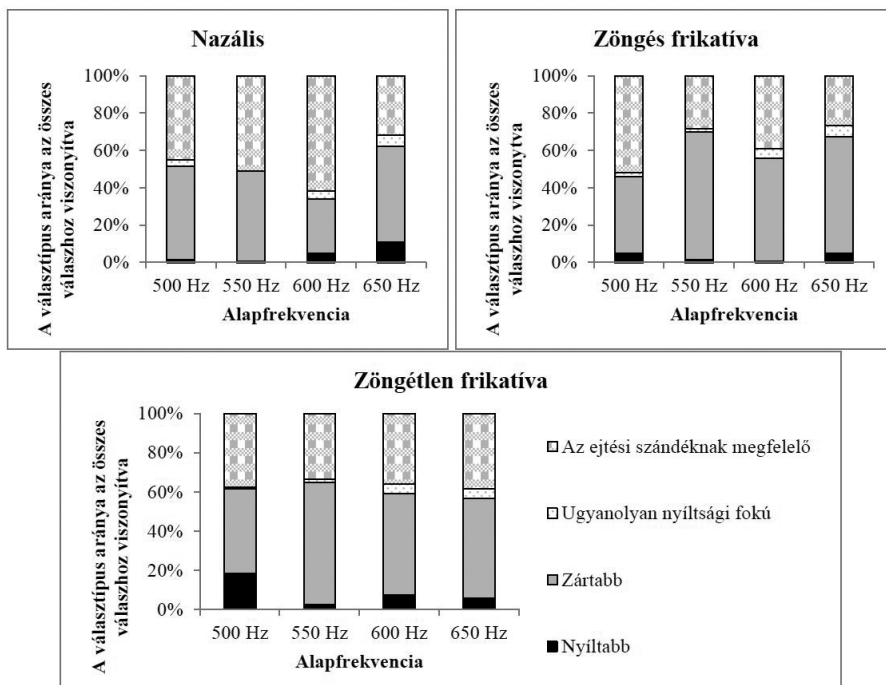
a legmagasabb alapfrekvencián nazális környezetben az /a:/, /u:/ és /e:/, zöngés frikatívák között az /a:/, /u:/ és /y:/, míg zöngétlen frikatívák között az /y:/ és az /o:/ azonosítása volt a legnagyobb arányban az ejtési szándéknak megfelelő. Annak a figyelembevételével, hogy kilenc választási lehetőség esetén annak az esélye, hogy az adatközlők véletlenszerűen választják az ejtési szándék szerinti magánhangzót, 11%, azt is megállapíthatjuk, hogy ezen a kísérletben vizsgált legmagasabb alapfrekvencián a nazális környezetben csak az /ø:/, míg a zöngés frikatívák között csak az /ɒ/ ejtési szándéknak megfelelő azonosítása csökkent a véletlen szintje alá, míg a zöngétlen frikatívák között egyetlen hang azonosítása sem. Bár az ejtési szándéktól eltérő minőségű magánhangzó-azonosításokkor adott válaszok szóródása az alapfrekvencia emelésével egyre nagyobbak bizonyult, a legmagasabb alapfrekvenciára kapott adatokból is látható módon a palatális és veláris magánhangzócsoportok még 650 Hz alapfrekvencián is egymástól elkülönülten jelentek meg a percepcióban.

4. táblázat: A kísérletben legmagasabb alapfrekvencián (650 Hz) tapasztalt azonosítási tendenciák a három mássalhangzó-környezetben

Nazális		Válasz (%)								
		/a:/	/ɒ/	/o:/	/u:/	/ø:/	/y:/	/ɛ:/	/e:/	/i:/
Az ejtési szándékozott magánhangzó	/a:/	63,3	23,3	3,3	10	0	0	0	0	0
	/ɒ/	23,3	30	33,3	13,3	0	0	0	0	0
	/o:/	16,7	13,3	40	30	0	0	0	0	0
	/u:/	3,3	3,3	36,7	56,7	0	0	0	0	0
	/ø:/	3,3	0	0	0	6,7	40	6,7	23,3	20
	/y:/	0	0	0	0	0	26,7	6,7	30	36,7
	/ɛ:/	0	0	0	0	6,7	10	30	43,3	10
	/e:/	0	0	0	0	0	3,3	3,3	50	43,3
	/i:/	0	0	0	0	3,3	3,3	3,3	73,3	16,7
Zöngés frikatíva		Válasz (%)								
		/a:/	/ɒ/	/o:/	/u:/	/ø:/	/y:/	/ɛ:/	/e:/	/i:/
Az ejtési szándékozott magánhangzó	/a:/	76,7	20	0	0	0	0	0	3,3	0
	/ɒ/	6,7	10	63,3	20	0	0	0	0	0
	/o:/	3,3	3,3	33,3	60	0	0	0	0	0
	/u:/	6,7	3,3	26,7	63,3	0	0	0	0	0
	/ø:/	0	0	0	0	20	60	3,3	10	6,7
	/y:/	0	0	0	0	26,7	50	3,3	13,3	6,7
	/ɛ:/	6,7	0	0	0	46,7	10	13,3	20	3,3
	/e:/	0	0	0	0	13,3	20	3,3	40	23,3
	/i:/	0	0	0	0	0	13,3	0	56,7	30

Zöngétlen frikativa		Válasz (%)								
		/a:/	/ɒ/	/o:/	/u:/	/ø:/	/y:/	/ɛ:/	/e:/	/i:/
Az ejteni szándékozott magánhangzó	/a:/	43,3	33,3	20	3,3	0	0	0	0	0
	/ɒ/	6,7	23,3	46,7	23,3	0	0	0	0	0
	/o:/	6,7	13,3	60	20	0	0	0	0	0
	/u:/	3,3	0	56,7	40	0	0	0	0	0
	/ø:/	3,3	0	0	0	20	60	0	10	6,7
	/y:/	0	0	0	0	16,7	73,3	0	6,7	3,3
	/ɛ:/	0	0	0	3,3	26,7	16,7	46,7	6,7	0
	/e:/	0	0	0	0	10	33,3	0	26,7	30
	/i:/	0	0	0	0	13,3	63,3	3,3	6,7	13,3

Ahhoz, hogy feltérképezzük az egyes alapfrekvenciákon jellemző, az ejtési szándéktól eltérő minőségű magánhangzó-azonosítások tendenciáit, ismét kiemeltük és összesítettük a második és harmadik nyíltsági fokú /o: ø: e: ɛ/ hangzókra kapott válaszokat az f_0 függvényében (18. ábra).



18. ábra: A második és harmadik nyíltsági fokú /e: ø: o:/ hangzókra az egyes alapfrekvenciákon kapott válaszok megoszlása a magánhangzók nyíltsági fokának függvényében

Bár az egyes hangkörnyezetekre kapott válaszok a magánhangzó észlelt nyíltsága szerint nem teljesen egységesek, jól látható, hogy minden középső vagy alsó nyelvvállású, azaz második vagy harmadik nyíltsági fokú magánhangzó esetében a magánhangzóknak az ejtési szándéknál zártabb azonosítása volt jellemző. Ezen túlmenően pedig azt a korábban (az átlagadatokon) tapasztalt tendenciát is megfigyelhetjük, hogy a nazális hangkörnyezetben a legmagasabb az ejtési szándéknak megfelelő azonosítás gyakorlatilag minden alaphangfokúán.

A mássalhangzó-környezeteket összehasonlítottuk a nem az ejtési szándék szerinti azonosítások átlagosan jellemző aránya szerint is, tehát aszerint, hogy egy adott környezetben összesen (az összes vizsgált alaphangfokúán együttesen) mekkora volt a nem az ejtési szándék szerinti azonosítások aránya. Ezek az értékek a következőképpen alakultak: nazálisok: 39%, zöngés frikatívák: 43%, zöngétlen frikatívák: 44%. A χ^2 -próba szerint a három hangkörnyezet e szerint a paraméter szerint nem egységes ($\chi^2 = 8,511$, $df = 2$, $p = 0,014$), amit a hibák számának (a 18. ábrán is látható) megoszlása alapján úgy értékelhetünk, hogy a nazális hangkörnyezet a zöngés és zöngétlen frikatíva kontextusnál magasabb arányban facilitálta az egyes magánhangzóknak az ejtési szándék szerinti azonosítását. (A nazális környezetben átlagosan 61%, a zöngés frikatívák között átlagosan 57%, míg a zöngétlen frikatívák között 56% volt az ejtési szándék szerinti azonosítás).

A hangkörnyezetek között kisebb különbség mutatkozik a nem az ejtési szándék szerint megjelölt válaszok magánhangzó-minőségek szerinti gyakorisági megoszlásában. Az 5. táblázatban összesített adatok szerint míg frikatíva környezetben a nem az ejtési szándék szerinti azonosításokkor a zárt /y:/ a leggyakoribb válasz, addig a nazálisok között ejtett magánhangzókra a leggyakrabban /i:/ válaszok érkeztek (mely szintén zárt, de az /y:/-vel szemben ajakréses magánhangzó).

5. táblázat: A nem az ejtési szándék szerinti azonosított magánhangzók esetében kapott válaszok gyakoriság szerinti sorrendje az egyes hangkörnyezetekben

	Nazális		Zöngés frikatíva		Zöngétlen frikatíva	
	Magánhangzó	Válasz aránya (%)	Magánhangzó	Válasz aránya (%)	Magánhangzó	Válasz aránya (%)
1.	/i:/	19,29	/y:/	22,38	/y:/	32,60
2.	/o:/	17,91	/o:/	15,49	/o:/	12,90
3.	/e:/	17,16	/ø:/	15,15	/ø:/	13,38
4.	/y:/	16,14	/u:/	14,80	/i:/	9,26
5.	/u:/	12,40	/e:/	12,22	/a:/	7,03
6.	/a:/	6,89	/i:/	8,09	/ɒ/	6,35
7.	/ɒ/	6,50	/ɒ/	7,06	/u:/	5,49
8.	/ø:/	1,99	/a:/	3,10	/e:/	5,32
9.	/ɛ/	1,77	/ɛ/	1,72	/ɛ/	0,69

3.3. Következtetések

Az imént bemutatott kísérletben azt a kérdést vizsgáltuk meg, hogy az előző vizsgálatból származó eredményeink mennyiben írhatók pusztán a vizsgálatban használt nazális mássalhangzó-környezet számlájára. Az előző vizsgálatnak a jelen kísérletben is ellenőrzött főbb eredményei a következők voltak. 1. Az alapprofrekvencia emelésével egyre csökkent az egyes magánhangzó-minőségek az ejtési szándéknak megfelelő azonosítása – ezzel párhuzamosan a válaszok egyre nagyobb szóródást mutattak. 2. A vizsgált közepesen magas alapprofrekvenciákon a zártabb és nyíltabb szomszédokkal is rendelkező énekelt magánhangzókat a nem az ejtési szándék szerinti azonosítások alkalmával a hallgatók inkább zártabbként azonosították. 3. A magas alapprofrekvencián legnagyobb arányban azonosított magánhangzók nem a legnyíltabb (legnagyobb nyíltsági fokú) magánhangzók voltak, hanem zárt, első nyíltsági fokú magánhangzók (az /i:/ és /y:/), valamint egy harmadik nyíltsági fokú magánhangzó (az /ɛ/).

A jelen fejezetben bemutatott kísérlet kiindulópontját a beszédben megvalósult nazalizált magánhangzók percepciójával kapcsolatos szakirodalom képezte. Eszerint bár a nazális mássalhangzó a koartikulációs nazalizációs hatáson keresztül megváltoztathatja a szomszédos magánhangzó hangszínét, a megjósolható helyen (nazális mássalhangzók környezetében) történt magánhangzó-nazalizációt a percepciói rendszer hatékonyan kompenzálja, így az azonosítás az obstruensek között ejtett nem nazalizált magánhangzókétól nem tér el jelentősen (KRAKOW et al. 1987). Az éneklés szakirodalmában a nazális hangszín megjelenését tekintve két, egymással ellentétes, vetélkedő nézet létezik: az egyik szerint az éneklésben nincs vagy csak nagyon csekély mértékű a nazális csatolás (VENNARD 1964; YANAGISAWA et al. 1990; YANAGISAWA et al. 1991), míg a másik szerint énekestől és magánhangzó-minőségtől függően változatos mértékben, de jelen van az orrüregi hatás az éneklésben (GRAMMING et al. 1993). A nazalitás megjelenése az éneklésben tehát vitatott kérdés, a nazalitás percepciói vetületéről ugyanakkor gyakorlatilag semmit sem tudunk az éneklés vonatkozásában, hiszen azt az említett vizsgálatok egyike sem ellenőrizte. KRAKOW és munkatársainak (1987) percepciói kísérletére alapozva a jelen vizsgálat elején azt feltételeztük, hogy a nazális és a frikatíva mássalhangzó-környezetek között nem találunk jelentősebb különbségeket az énekelt magánhangzók percepciójában.

Az eredmények csak részben igazolták az elvárásainkat. Általánosan ugyanis azt láttuk, hogy nem volt jelentősebb eltérés az egyes hangkörnyezetek között, apróbb különbségeket azonban több ponton is találtunk, és ezek egy része éppen a nazális hangkörnyezet pozitív hatását mutatta az énekelt magánhangzók észlelésében a frikatívákkal szemben.

1. Az adatok szerint az adatközlők az alapprofrekvencia emelésével egyre kisebb számban azonosították a magánhangzókat az ejtési szándék szerint kontextustól függetlenül, úgy, hogy az egyre magasabb alapprofrekvenciák a válaszok egyre nagyobb szóródását, tehát a magánhangzó-minőségek egyre fokozódó percepciói bizonytalanságát mutatták. Ezzel együtt azt

láttuk, hogy a nazális hangkörnyezetben ejtett magánhangzók a vizsgált köztes (500 Hz és 550 Hz alapfrekvenciájú) énekelt hangmagasságokon valamivel nagyobb arányban idézték elő az ejtési szándék szerinti azonosítást. A magánhangzó-azonosítást a három mássalhangzó-környezet szerint összesítve azt találtuk, hogy a nazális hangkörnyezetben valamivel magasabb volt az ejtési szándék szerinti azonosítások aránya. Kiemelendő, hogy a beszédben KRAKOW és munkatársainak (1987) eredményeihez hasonlóan azt találtuk, hogy a nazális és obstruens környezetben nincs különbség a magánhangzók az ejtési szándék szerinti azonosításának arányaiban.

2. A második és harmadik nyíltsági fokú (zártabb és nyíltabb szomszédokkal egyaránt rendelkező) /e e: ø: o:/ magánhangzók vizsgálatok az láttuk, hogy ha ezeket a magánhangzókat nem az ejtési szándék szerinti azonosítják a hallgatók az éneklésben, akkor inkább zártabbként ismerik fel őket, ugyancsak kontextustól függetlenül. Emellett egyik kontextusban sem láttunk szoros összefüggést az alapfrekvencia és a nem az ejtési szándék szerinti azonosításokat jellemző tendenciák között. Úgy véljük, hogy a jelen kísérlet bizonyítékkal szolgál arra, hogy a korábbi szakirodalomban eddig nem ismertetett, az észlelést az énekelt magánhangzók esetében jellemző tendencia (mely szerint a zártabb és nyíltabb szomszédokkal is rendelkező magánhangzókat közepesen magas alapfrekvenciákon a nem az ejtési szándék szerinti azonosítások esetében inkább zártabbként azonosítják a hallgatók) a magas alapfrekvencián énekelt magánhangzók azonosításának egy igen fontos jellemzője, ami a beszédpercepció és a produkció közti nem egyszerű és lineáris összefüggésekről árulkodik.

A korábbi percepció vizsgálatok nagy hangsúlyt fektettek azon magánhangzók elemzésére, melyeket – első nyíltsági fokuk lévén – a hallgatók a nyíltsági foknak nem az ejtési szerinti azonosítások nem is azonosíthatnak zártabbként, csak nyíltabbként (HOLLIER et al. 2000), illetve nem adtak pontos, számszerű információkat arról, hogy mire alapozták azt a feltételezésüket, mely szerint az alapfrekvencia emelését az egyes magánhangzók nyíltabbként azonosítása jellemezné (SCOTTO DI CARLO–GERMAIN 1985). Mindemellett viszont arra is utaltak, hogy az egyes zárt, azaz első nyíltsági fokú magánhangzók (elsősorban az /i:/) azonosítása kiemelkedően magas volt még az extrém alapfrekvenciákon is (SCOTTO DI CARLO–GERMAIN 1985). Ennek megfelelően e korábbi tanulmányok, bár megfogalmazták a produkció és a percepció általuk feltételezett egyszerű, lineáris összefüggéseit – mely szerint a produkciós nyitódás egyúttal nyíltabb észleletet is eredményez – valójában nem alapozták meg maradéktalanul e feltételezésüket.

A korábbiakkal szemben a 2. és 3. fejezetben bemutatott vizsgálataink jól kontrollált, szisztematikus és a magánhangzórendszer sajátosságainak figyelembevételével elvégzett elemzéseknek tekinthetők, melyek bizonyítékkal szolgáltattak arra, hogy a produkció és a percepció közti kapcsolat az énekelt magánhangzók azonosításában sem tekinthető egyszerű, lineáris összefüggésnek (sőt azt az egyszerűbb lineáris összefüggések alapján egyáltalán nem értelmezhető tendenciák jellemzik). Mindez erős bizonyítéknak tűnik a beszédpercepció szakirodalmában már több évtizede ismeretes, a produkció és a percepció bonyolultabb összefüggéseiről szóló ismeretekre is (vö. SRULOVICZ–GOLDSTEIN 1983; DELGUTTE–KIANG 1984;

CHISTOVICH–LUBLINSKAJA 1979; CARLSON et al. 1970; TRAUNMÜLLER 1981 és a jelen értekezés 1.1.2.2. alfejezete).

3. Az előzőekkel is összefüggő módon a jelen vizsgálatban elemzett legmagasabb alapfrekvencián nyíltabb és zártabb minőségű hangzók is voltak azok között, melyeket nagyobb arányban azonosítottak az ejtési szándék szerint a tesztben részt vevők (a nazálisok esetében: /a:/, /u:/ és /e:/, a zöngés frikatívák esetében: /a:/, /u:/ és /y:/, a zöngétlen frikatívák esetében: /y:/ és /o:/). E szempont szerint a fonetikailag zöngés nazális, valamint a nyelvilleg is zöngés frikatíva egymáshoz valamelyest közelebb állónak látszik, mint a más szempontok szerinti összehasonlításokban.

A vizsgálatban a nazális hatást mint az ejtési szándék szerinti azonosítást akadályozó paramétert teszteltük énekes adatközlőnk produkciójában. Azt a feltételezést, hogy a nazális hangkörnyezet valóban ilyen hatást fejtene ki a magánhangzók azonosítására, eredményeink alapján egyértelműen elutasíthatjuk. Ráadásul a nazális mássalhangzók környezetében ejtett magánhangzók nemcsak, hogy nem mutattak alacsonyabb azonosítási arányokat, hanem sokszor éppen hogy magasabb arányban voltak az ejtési szándék szerint azonosíthatók, mint a frikatívák között ejtett, magas alapfrekvenciákon énekelt magánhangzók. Ennek értelmében tehát megfogalmazhatjuk azt a sejtést is, hogy a nazális mássalhangzók nem pusztán a „zengető” hatás miatt (vö. KERÉNYI 1959; VÁRADI 2010) lehetnek kitüntetett szerepűek az éneklésben. Ha ugyanis a magánhangzók percepciók elkülönítése a nazális mássalhangzók környezetében bizonyos esetekben (bizonyos alapfrekvenciák esetében) jobban biztosított, akkor az énekesek (és énektanárok) számára a nazálisok használata által megerősítést nyerhet az eddigi adataink alapján inkább csak illuzionisztikusnak tűnő, de az énektechnika szempontjából valószínűleg nagyon fontos élmény és érzemény, mely szerint az egyes magánhangzók minősége egymástól elkülöníthető a magas alapfrekvenciájú éneklésben. Mivel e feltételezés tesztelésére előbb vizsgálni kellene az énekelt magánhangzók nazalizációjának lehetséges mértékét, majd ennek ismeretében a különböző realisztikus mértékben nazalizált magánhangzók észleletének alakulását is, ezért a felvetés ellenőrzése messze túlmutat a jelen kutatás keretein, és e helyütt csak feltételezés maradhat.

Fontos kiemelnünk, hogy a jelen kísérletben részt vevő énekes üregi rezonanciái, illetve a nazalitással kapcsolatos produkciós stratégiája (tehát a nazális csatolás az énekes egyéni fiziológiai és produkciós jellemzői tekintetében) feltehetően sajátos, egyénien érvényesültek. Ez azt jelenti, hogy énekes adatközlőnk nazalizációs stratégiái nem feltétlenül képezik le az operai énektechnikában elvárt vagy tradicionálisnak tekinthető stratégiát; miközben azt sem szabad elfelejtenünk, hogy egyes korábbi vizsgálatok sugalmazásai szerint nem is teljesen valószínű, hogy ilyesfajta egységes nazalizációs stratégia feltételezhető az éneklésben (vö. GRAMMING et al. 1993). A jelen vizsgálatnak tehát nem volt, de nem is lehetett célja a nazalitás megjelenésével kapcsolatos általános érvényű produkciós stratégiákat feltérképezni – kiváltképp akkor, ha elfogadjuk az imént említett tanulmány javaslatát a nagy egyéni variabilitásról. Célunk ezzel szemben csak az lehetett – és ennek a kísérlet, úgy véljük, maradéktalanul eleget is tett –, hogy bemutassuk, hogy az előző kísérletünkben, az ugyanezen

énekes adatközlő produkálta magánhangzók azonosításában talált, a produkció és percepció nem egyszerű, lineáris összefüggéseit mutató eredmények egyértelműen nem a nazális hang-környezet, hanem a magas alaphangfrekvenciájú éneklés mint produkciós mód sajátosságainak következtében alakultak a tapasztaltak szerint.

A megelőző és a jelen kísérletünk eredményei alapján, úgy véljük, megfogalmazható az a kijelentés, mely szerint nincs okunk feltételezni a produkció és percepció közti, a korábbi vizsgálatokban sugallt nagyon leegyszerűsítő összefüggést, mely szerint az éneklés során az alaphangfrekvencia emelésével egyre fokozódó produkciós nyitódást (azaz az állkapocsnyitás szögének növelését, illetve az F_1 emelését) egyértelműen percepciós nyitódás is kísérné. Az adatok alapján sokkal inkább azt mondhatjuk, hogy az alaphangfrekvencia emelését kísérő produkciós változások hatására a magánhangzók percepciós minősége is egyre bizonytalanabb lesz az éneklésben (eddiggi adataink alapján a 650 Hz közepesen magas alaphangfrekvenciáig). Az egyes hangzóminőségeknek az alaphangfrekvencia emelésével egyre fokozódó percepciós változásairól ezen felül úgy láttuk, hogy azok graduális változást jelentenek – az alaphangfrekvencia emelésével egyre jobban megváltozó akusztikai szerkezetű magánhangzók tehát nem kategorikusan, az egyik magánhangzó-kategóriáról a másikra ugrásként csapódnak le az észleletben.

A következő kísérletünkben az éneklés szakirodalmában azt a gyakran empirikus bizonyítékok nélkül is elfogadott feltételezést kívánjuk megvizsgálni, mely szerint a koartikuláció akusztikai vetületének, azaz a formánsátmeneteknek köszönhetően a magas alaphangfrekvencián énekelt magánhangzók a beszédhangokhoz hasonlóan magasabb arányban azonosíthatók az ejtési szándék szerint, mint az izoláltan ejtett magánhangzók. Mivel eddiggi akusztikai eredményeink azt mutatták, hogy az énekelt magánhangzók akusztikai szerkezete megváltozik az alaphangfrekvencia emelésével, percepciós eredményeink szerint pedig ez a változás elbizonytalanítja a magánhangzók észlelt minőségét, úgy véljük, hogy a vizsgálandó feltételezés bonyolultabb annál, mintsem hogy a beszédpercepció szakirodalmából ellenőrzés, azaz szisztematikus és jól kontrollált kísérletekből nyert empirikus bizonyítékok nélkül adaptálható legyen.

4. A mássalhangzó-környezet hatása az énekelt magánhangzók észlelésére: azonosítás mássalhangzó-környezetben, izolált ejtésben, valamint a zöngéindítás mint akusztikai kulcs hiánya esetén

Mint azt a korábbi, valamint a jelen értekezés *2. fejezetében* bemutatott produkciós vizsgálatokban láttuk (továbbá amint azt a *6. fejezetben* bemutatandó kísérletünk is demonstrálni fogja), az énekelt magánhangzók produkciója igen erősen függ az alapfrekvenciától, olyan módon, hogy a magánhangzók artikulációs és akusztikai megvalósítása az alapfrekvencia emelésével egyre inkább eltér a magánhangzók beszédbeli megvalósításától. Amint azt az eddig bemutatott szakirodalmi leírások és saját észlelési kísérleteink bizonyítják, e produkciós változás – természetesen – a magánhangzók észleleti vetületét is erősen befolyásolja, általánosságban olyan módon, hogy az alapfrekvencia emelésével egyre csökken az ejtési szándék szerinti magánhangzó-azonosítások aránya. Mégis, mindennapi tapasztalataink azt mutatják, hogy a nem laboratóriumi körülmények között rögzített, nem mesterségesen generált, értelmetlen hangsorokból álló stimulusok, hanem az adott nyelven értelmes szavakká, majd értelmes szöveggé formálódó beszédhangok éneklése során jó arányban lehetséges a magánhangzók az ejtési szándék szerinti azonosítása (hiszen többé-kevésbé megértjük a dalok és operaáriák szövegét). Jogosan merül fel tehát a kérdés, hogy mennyiben segítheti a mássalhangzó-környezet és nagyobb kontextusban az értelmes szavakba és mondatokba ágyazottság az egyes magánhangzók azonosítását.

A kontextus szerepe az énekelt magánhangzók észlelésében, ahogy a beszédpercepcióban is, valójában kétarcú probléma. Az énekelt magánhangzók észlelését eddig vizsgáló szakirodalomban azonban az látható, hogy e két, élesen elkülönítendő oldal nem válik szét hangsúlyosan (erről már a szakirodalmat bemutató *1.3.3.2. alfejezetben* is értekeztünk részletesebben, de a későbbiekben még több ponton utalunk rá). A kontextus kérdésének egyik aspektusa a pusztán a beszédmegértés legalsóbb szintjeire támaszkodó bottom-up folyamatok aktivációjával kapcsolatos beszédészlelési probléma: miként és mennyiben segíti a magánhangzó-azonosítást a magánhangzók a szomszédos mássalhangzókkal mind időben, mind térben átfedő, együttes ejtésének, azaz a koartikulációnak az akusztikai vetülete, a koartikulációs formánsátmenet? A kérdés másik aspektusa azonban nemcsak az akusztikai elemzést végző bottom-up folyamatokra támaszkodó, de a beszédmegértési folyamat felsőbb szintjeinek (tehát a top-down folyamatoknak az) aktivációjához is köthető észlelési probléma. Eszerint ha a magánhangzó-felismerés értelmes, valamint esetleg szintaktikailag és szemantikailag is

beágyazott szavakban zajlik, a lehetséges – tehát a felismerés kimeneteként szóba jöhető – magánhangzóválaszok száma és minősége jóval korlátozottabb, mint a nem értelmes, izolált szavak (vagy izolált hangzók) esetén. Ennek megfelelően ezekben a felismerési helyzetekben valójában egyáltalán nem szükséges a maradéktalan akusztikai dekódolás az egyes magánhangzók minőségének felismeréséhez – az akusztikailag nem beazonosítható magánhangzókat képesek vagyunk „kitalálni” a részben vagy egészben dekódolt kontextus alapján.

A mássalhangzó-környezet mint a magánhangzóknek az ejtési szándék szerinti azonosítását segíthető jelenség tehát két, egymástól eltérő aspektus mentén is vizsgálandó jelenség. Az ebben a fejezetben bemutatott vizsgálat a két aspektus közül az elsőre fókuszál, a kísérlet fő célja annak a kérdésnek a megválaszolása, hogy mennyiben segíti a mássalhangzó-környezetet a formánsátmenetek révén a magas alapfrekvencián énekelt magánhangzóknek az ejtési szándék szerinti azonosítását. A kérdés másik aspektusát, egészen pontosan a két aspektus közti különbségek módszeres vizsgálatát, azaz a formánsátmenetek és az értelmes szavakba ágyazottság hatásának összevetését a következő fejezetben bemutatásra kerülő kísérletünkben teszteljük. A jelen kísérletben ugyanakkor még egy további kérdésre is választ keresünk, ugyanis az énekelt magánhangzók esetében is meg kívánjuk vizsgálni azt a pszichoakusztikai szakirodalomban leírt jelenséget, mely szerint a hangszín azonosításában kiemelt szerepe van a hangindításnak.²²

GOTTFRIED és CHEW (1986) kísérletében mássalhangzó-környezetben ejtett magánhangzók azonosítását vetette össze a kontextusból kivágott magánhangzó-tisztafázisok észlelésével. Bár éppen emiatt a kísérlet eredményei nem vonatkoztathatók az izoláltan ejtett magánhangzók és a mássalhangzó-környezetben ejtett magánhangzók észlelése közti különbségekre, a kísérleti elrendezés felhívja a figyelmet egy másik izgalmas és tesztelendő kérdésre. Egyes pszichoakusztikai kísérletek eredményei szerint ugyanis egy adott alapfrekvencián megszólaló hang hangszínének azonosításában, azaz a hangszerek keltette hangok világában a hangot kiadó hangszer felismerésében kiemelt szerepű akusztikai kulcs a hangindítás, ezen információ hiányában ugyanis jelentős mértékben csökken a hangszerek felismerésének sikeressége (BERGER 1964; SALDANHA–CORSO 1964). Felmerülhet tehát a kérdés, hogy vajon a hangadás kezdetének kiemelt jelentősége van-e a beszédhangok világában is, ahol az egyes hangszínbeli különbségeket az egyes magánhangzó-minőségek különbségeinek feleltethetjük meg. Bár GOTTFRIED és CHEW (1986) a magánhangzók tiszta fázisának tesztelésével gyakorlatilag megvizsgálja a kérdés egyik oldalát (ez volna a hangindítás eliminálásával kapott stimulusanyag), és összeveti azt a kontextusban ejtett magánhangzók azonosításával, nyitott kérdés marad, hogy milyen eredményekre vezetne a magánhangzók tiszta fázisára kapott észlelési adatok összevetése az izolációban ejtett énekelt hangzók azonosításával. E kérdés vizsgálata képezi tehát a jelen fejezetben bemutatásra kerülő kísérlet második célját.

²² Az ötlet, hogy a zöngéindítás mint akusztikus kulcs magánhangzók azonosítására hatással lehet, 2012-ben, egy Svédországban megrendezett nyári egyetemen merült fel először egy Svante GRANQVIST tartotta workshop után, az előadóval beszélgetve. Bár az akkor felmerült ötletre épülő vizsgálatban Svante GRANQVIST már nem kívánt részt venni, az említett beszélgetés így is nagy lendületet adott a kísérletek lefolytatásához, amiért azóta is hálás vagyok.

Mielőtt rátérnénk a kísérlet anyagának és módszertanának ismertetésére, fontosnak tartjuk emlékeztetőül röviden felidézni az 1.2.2. *alfejezet*ben tárgyalt, a szomszédos mássalhangzók hatását az énekelt magánhangzók azonosításában vizsgáló szakirodalom problematikus pontjait, melyek alapján úgy véljük, hogy a mássalhangzó-környezet hatása az énekelt magánhangzók észlelésében nem tekinthető sem alaposan körüljárt, sem pedig megnyugtatóan megválaszolt kérdésnek.

Azon feltételezés tudományos igazolására, hogy a mássalhangzó-környezet megléte vagy hiánya hatással van-e a magas alapfrekvencián (szopránéneklésben) ejtett vokálisok azonosítására, eddig összesen egyetlen kutatás született. (Egy további vizsgálat, a már említett GOTTFRIED és CHEW (1986) szerzőpáros munkája mélyebb alaphangokon, egy férfi hangkategóriában, egy tenor énekes ejtésében vizsgálta a kérdést.) SMITH és SCOTT (1980) a magas alapfrekvencián énekelt magánhangzók azonosítását *bVd* hordozó hangsorban vizsgálta egy szoprán énekes ejtésében. Percepciós adataik a szerzők értelmezése szerint az operaéneklésre nézve is megerősítik azt a beszédben tapasztalható tendenciát, mely szerint a szomszédos mássalhangzók segítik a magánhangzók felismerését. Adataikat a szerzők azonban nem szigorúan kontrollált körülmények között nyerték, a módszertan több okból is nehezen értelmezhető eredményekhez vezetett.

Egyrészt a magánhangzót közrefogó mássalhangzók képzéshelye eltérő volt, így valójában nem lehet biztosan tudni, hogy mely mássalhangzó befolyását tesztelték a szerzők, sem pedig azt, hogy van-e különbség a magánhangzót megelőző és követő hangátmenetek hatása között.

Másrészt az sem egyértelmű, hogy milyen módon kontrollálták a szerzők az énekes gégejének függőleges helyzetét, ami pedig – állításuk szerint – szintén fontos befolyásoló tényezőnek bizonyult. A kísérlet leírásából úgy sejthető, hogy a „gégehelyzet” pusztán csak egy, a szerzők egyéni percepciós megítélésén alapuló paraméter, tehát valamilyen percepciós alapú, de nem pontosan definiált kategória volt.

Egy harmadik probléma, hogy a szerzők feltételezése szerint az $F_1 : f_0$ formánshangolás mint az operai énekemód sajátossága módosította a beszédben jellemző akusztikai szerkezetet, és így csökkentette az ejtési szándék szerinti azonosíthatóságot; ezzel találták a szerzők magyarázhatónak azt, hogy az izolált ejtésű magánhangzók az ejtési szándék szerinti azonosításának aránya alacsonyabb volt az operai módban a „nem rendes” énekprodukciónhoz képest. Ha viszont ennek ellenére a mássalhangzó-környezetben mégis magasabb a magánhangzók ejtési szándék szerinti azonosítása, mint izolált ejtésben, és ezt a szerzők kizárólag a mássalhangzó-környezet számlájára írják, akkor az azt jelenti, hogy a szerzők szerint valójában a formánsstratégia, azaz az $F_1 : f_0$ hangolás érvényesült a kontextus (tehát a mássalhangzó-környezet megléte vagy hiánya) függvényében. Ilyen értelemben tehát a szerzők valójában azt a feltételezést implikálják tanulmányukban, hogy a két kondícióban megvalósult magánhangzók akusztikai szerkezete eltérő lehetett, és ez tette lehetővé azt, hogy a hallgatók különböző minőségű magánhangzóként ismerhették fel azokat.²³

²³ A vizsgálat eredményei a 2013-ban megrendezett Beszédkutatás konferencián hangzottak el először, amikor is a közönség több tagja is kérdéseket, megjegyzéseket fűzött az előadáshoz. Ezek között Törkenczy Miklós azt a kérdést tette fel, hogy hogyan lehet értelmezni az ejtési szándék szerinti hangzóazonosítást támogató körül-

Egy negyedik és a vizsgálat eredményeinek értelmezését legjobban megnehezítő probléma pedig a kísérlet alapvető felépítésének, módszertanának egy meghatározó sajátossága. Míg a teszt során a résztvevők a CVC kondícióban négy lehetőségből, négy értelmes szó közül választhattak, addig az izolált ejtésű hangok esetében a választási lehetőségek egyáltalán nem korlátozódtak arra a négy hangzóra, amelyet az értelmes szavak megengedtek. Az izolált ejtésben ugyanis a válaszlehetőségek számát csak az angol nyelv magánhangzókészlete korlátozta, mely legalább tíz lehetséges magánhangzóválaszt hagyott. Ennek fényében tehát SMITH és SCOTT (1980) a magánhangzóknak a mássalhangzó-környezet meglététől függő azonosítását valójában két jelentősen eltérő kondíció összevetésével igyekezett vizsgálni, melyekhez ráadásul két jelentősen eltérő percepciós működést kapcsolhatunk. Az egyik vizsgált kondícióban (CVC) az értelmes szavak révén a nyelv magasabb szintjei, a morfológia, illetve a szemantika is befolyásolta az azonosítás tendenciáit, hiszen ezek korlátozták a válaszlehetőségeket és a válaszlehetőségek számát a válaszként szóba jöhető létező szavak (azaz a top-down percepciós folyamatok) révén. A másik, az izolált ejtésű hangzókat tartalmazó kondícióban ezzel szemben csak az alacsonyabb szintű, elemi hallási észlelésen és az akusztikai elemzésen alapuló (és csak az akusztikai kulcsokra támaszkodó, bottom-up) beszédpercepciós folyamatok léphettek működésbe úgy, hogy a lehetséges válaszok számát az angol nyelv magánhangzókészletén kívül semmi sem korlátozta. SMITH és SCOTT vizsgálatában tehát nem választotta szét a mássalhangzó-környezet két szigorúan elkülönítendő aspektusát, és így bár az eredmények elsősorban a beszédmegértés magasabb folyamatainak aktivációjából is származó percepciós működésekről tanúskodtak, a szerzők a tapasztalt pozitív hatásokat kizárólag a formánsátmenetek mint akusztikus kulcsok javára írták.

Mindezek értelmében SMITH és SCOTT (1980) kísérlete valójában nem ad egyértelmű választ a mássalhangzó-környezet hatásának kérdéseire, ráadásul a vizsgálati elrendezés a kondíciók több ponton eltérő feltételei miatt a vizsgálati eredmények több, egymással interakcióban lévő zavaró (*confound*) hatás eredményeként álltak elő (melyek támogatták, erősítették a különbségeket a mássalhangzó-környezet hatását reprezentálni kívánó kondíció és az izolált ejtést reprezentáló kondíció között).

A jelen fejezetben ismertetett kísérlet feltételezése (az 1.3. *alfejezet*ben is megfogalmazott hipotézissel összhangban) a következő. Abból kiindulva, hogy az f_0 emelésével a magánhangzók spektrális alul-mintavételezettsége fokozódik, valamint az $F_1 : f_0$ hangolás következtében az f_0 emelésével az énekelt magánhangzók formánsszerkezete is egyre inkább megváltozik, azt feltételezzük, hogy a magas alapfrekvenciákon énekelt magánhangzók azonosítása is egyre kevésbé történhet az ejtési szándék szerint. Mivel azonban az alapfrekvencia emelésével egyre fokozódó változás magukat a magánhangzó-megvalósulásokat (tehát a produkciót), nem pedig a percepciót érintő folyamat, azt is feltételezzük, hogy sem a mássalhangzó-

ményeket annak figyelembevételével, hogy feltételezhetően a magas alapfrekvenciákon megváltozik a magánhangzók akusztikai szerkezete. A válasz végiggondolása hozzásegített, hogy a szakirodalomban található ellentmondást világosabban láttassam, kutatásom vizsgálati kérdését pedig egyértelműbben ragadjam meg, a kérdést ezúton is köszönöm.

környezetben rejlő dinamikus információ, sem pedig a zöngéindítás nem segítheti a percepció működéseket abban, hogy ezeket a folyamatokat kompenzálják, és segítsék a hallgatót a magas alapfrekvencián énekelt, az f_0 emelésével egyre bizonytalanabb (azaz a prototipikus kategóriák közé eső) minőségű magánhangzók az ejtési szándék szerinti azonosításában. Az alacsonyabb alapfrekvenciák (azaz a beszédbeli megvalósuláshoz hasonlóbb akusztikai minőségű énekelt magánhangzók) esetében ugyanakkor azt feltételezzük, hogy a mássalhangzó-környezet bizonyos mértékű pozitív hatása kimutatható lesz a közrezárt magánhangzónak az ejtési szándék szerinti azonosításában. A zöngéindítás ennél valamelyest enyhébb, de mégis pozitív hatását feltételezzük, mely (az alacsonyabb alapfrekvenciákon) a zöngéindítás nélkül prezentált hangok azonosításához képest valamivel magasabb, az ejtési szándék szerinti azonosítási arányokat eredményez.

4.1. Anyag, módszer és kísérleti személyek

A percepció vizsgálatához egy professzionális szinten éneklő szoprán énekesnő (életkora 27 év) énekprodukciónak rögzítettük tesztanyagként. Az énekes a magyar nyelv három, akusztikai tekintetben egymástól legtávolabb eső magánhangzóját /a: i: u:/ ejtette izoláltan („V” kondíció) és *bVb* hangkörnyezetben („CVC” kondíció), hat énekelt alapfrekvencián (ezeket a 6. táblázatban összesítettük) és beszélve (a beszédanyag átlagos alapfrekvenciája 191 Hz volt). A vizsgálat újítása eddigi produkciók és percepciók kísérleteinkhez képest az, hogy ez alkalommal a temperált zenei skálák figyelembevételével választottuk meg a vizsgált alapfrekvenciákat, olyan módon, hogy azok (a hangmagasságérzet szerint) egymástól egyenlő távolságokra helyezkedjenek el, valamint gyakorlatilag teljesen lefedjék a szoprán énekesek által használt alapfrekvencia-tartományt. A felvételeket csendesített szobában rögzítettük 44,1 kHz-es mintavételezéssel, 16 biten.

6. táblázat: A kísérletben használt zenei hangok és azok frekvenciaértéke

Zenei hang	Alapfrekvencia (Hz)
f	175
h	247
f'	349
h'	494
f''	698
h''	988

A teszthangsorokon kívül disztraktor hangsorokat is rögzítettünk ugyanazon énekesssel. A disztraktorok további magyar köznyelvi magánhangzókat /ɒ ε e: o: ø: y:/ tartalmaztak *bVb*

hangkörnyezetben és izolált ejtésben, valamint beszélve, illetve az f' , h' , f'' zenei alaphangokon énekelve. (Ezen stimulusok percepcióis eredményeit nem elemeztük.)

A hordozó hangsor (azaz a *b-b* váz) megválasztását a következők motiválták. A mássalhangzó-környezet hatásának szigorú kontrollálása érdekében olyan CVC hordozó hangsort kellett használnunk, melyben a két mássalhangzó azonos, így az adataink egyértelműen az adott képzéshelyű mássalhangzóra vonatkoztathatóak. Ahhoz azonban, hogy az eredményeinket a SMITH és SCOTT (1980) kísérletében látottakkal is összehasonlíthatóvá tegyük, az általuk használt *bVd* hangsor egyik mássalhangzóját kellett megtartanunk. Korábbi kutatási tapasztalataink szerint a CVC hangsor első mássalhangzója gondosabban artikulált az éneklésben, mint az utolsó, ezért ennek a nagyobb befolyását feltételeztük SMITH és SCOTT (1980) eredményeire, és a *bVd* hangsor első mássalhangzójának a megtartása mellett döntöttünk, amelyet egyúttal a kísérlet nyelvi ingereinek végére is helyeztünk (így kaptuk tehát a *b-b* vázat).

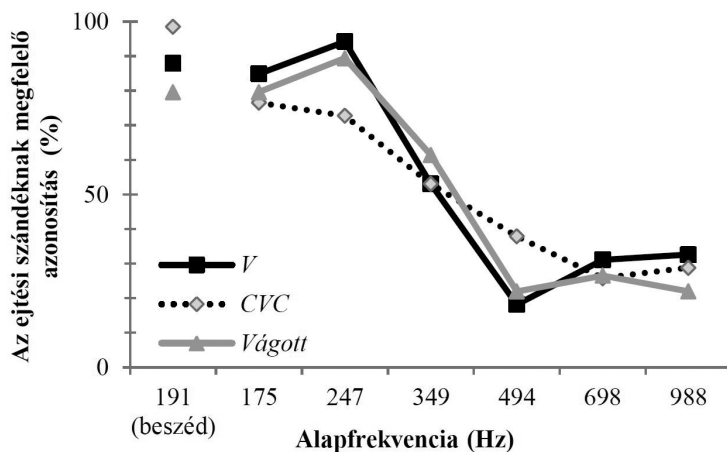
A magánhangzó zöngéindulásának hatását egy harmadik kondíció előállításával teszteltük: egy hangedítáló program, a WaveSurfer (SJÖLANDER–BESKOW 2000) segítségével („fokozatos beúszás” hatás alkalmazásával) eltávolítottuk az izoláltan ejtett hangok elejét („Vágott” kondíció). A beúsztatást 30 ms-os időtartamon, logaritmikus függvény segítségével valósítottuk meg.

A hanganyagokat egy, a Praat programban (BOERSMA–WEENINK 2009) írt és futtatott percepcióis tesztben vizsgáltuk 22 ép hallású felnőtt részvételével (21 és 35 év közöttiek, átlagos életkoruk 29 év). A teszt (3 kondíció \times 3 magánhangzó \times 7 alaphangmagasság \Rightarrow 63 stimulust és 30 disztraktort tartalmazott. Minden tesztelő minden stimulust kétszer hallott, véletlenszerű sorrendben, így egy ülésben összesen 156 hangminta hangzott el. A teszt előtt a kísérleti személyek azt az utasítást kapták, hogy azonosítsák az önmagában vagy hangsorban elhangzó magánhangzót, és döntésüket a képernyőn megjelenő kilenc magánhangzó egyikének megjelölésével rögzítsék. Ehhez a képernyőn válaszként a magyar hosszú (vagy hosszan ejtendő) magánhangzókat /ɒ a: ε e: i: o: ø: u: y:/ jelenítettük meg ortografikus alakban. A tesztelők a stimulusokat binaurálisan, fejhallgatón keresztül hallották, és a számítógéphez csatlakoztatott egérrel kattinthattak a kiválasztott magánhangzóra. A teszt adatközlőnként átlagosan 10-15 percet vett igénybe, mivel azonban a stimulusok sorrendje minden tesztben eltérő volt, nem volt szükséges figyelembe vennünk a fáradás következtében esetlegesen fellépő teszt-hatást.

A válaszokat ismét tévesztési mátrixokban összesítettük minden stimulus-válasz párt alaphangmagasságként rendezve. Az adatok statisztikai elemzését az R programmal végeztük el (R Core Team 2013). Az elemzés során előbb Shapiro–Wilk-próbával elemeztük a normális eloszlás kritériumát, melynek teljesülése esetén parametrikus tesztet alkalmaztunk (ANOVA), ellenkező esetben pedig nemparametrikus tesztet (χ^2 -próba, Mann–Whitney *U*-próba). A korrelációs számításokat a Pearson-féle korrelációelemzéssel végeztük.

4.2. Eredmények

A 19. ábra a három vizsgált magánhangzónak az ejtési szándék szerinti azonosításának százalékos arányát (y tengely) mutatja az alaphangmagasság függvényében (x tengely), a három kondíció szerint szétválasztva, de mindhárom magánhangzó-minőséget összesítve. Az adatok megjelenítése azt sugallja, hogy a vokálisoknak az ejtés szerinti azonosítása csökken az f_0 emelésével. A kéttényezős ANOVA szerint, melyben az „ f_0 ” és a „kondíció” hatását teszteltük, ez a benyomás helyes, és az f_0 paraméter valóban hatással van az ejtési szándék szerinti azonosítási arányok alakulására [$F(1) = 33,33, p < 0,001$]. Az összes lehetséges pár összehasonlítása (Bonferroni-korrekcióval) azt is felfedte, hogy az egymást alaphangmagasság szerint követő énekelt stimulusok között egészen a h' zenei hangig (494 Hz) mutatkozik páronként szignifikáns eltérés ($p < 0,02$), mely alaphangtól felfelé azonban a különbségek már nem ilyen jelentősek. A Pearson-féle korrelációs tesztek azt mutatták, hogy az „ f_0 ” és az „ejtési szándéknak megfelelő azonosítás aránya” változók kapcsolatát viszonylag erős negatív korreláció jellemzi. Amennyiben a három magánhangzóra kapott adatokat egy csoportban vizsgáljuk, erős összefüggést találunk: $r = -0,631, p < 0,001$. Ha pedig szétválasztjuk az adatokat a három kondíció mentén, némileg eltérő erősségű, de minden esetben szignifikáns összefüggéseket látunk: „V”: $r = -0,590, p = 0,005$; „CVC”: $r = -0,609, p = 0,003$; „Vágott”: $r = -0,703, p < 0,001$.



19. ábra: A három vizsgált magánhangzó (a: i: u/) azonosításának összesített százalékos aránya az alaphang függvényében a három vizsgálati kondícióban

A kéttényezős varianciaanalízis (ANOVA) nem igazolta, hogy a „kondíció” is hatással lenne az észlelési eredményekre (sem önmagában, sem pedig az „ f_0 ” változóval interakcióban).

Ráadásul a 19. ábrát szemlélve az is kitűnik, hogy a „kondíció” és az „ejtési szándék szerinti azonosítás” közti összefüggés valójában nem is volt konzisztens. Eszerint nemcsak, hogy nem teljesül az a várakozás, mely szerint a magánhangzókat nagyobb arányban lehetséges az ejtési szándék szerint azonosítani mássalhangzó-környezetben, de az is igazolódni látszik, hogy néha (például a *h* zenei alaphangon) kifejezetten alacsonyabb arányú a CVC szekvenciában elhangzó vokálisoknak az ejtési szándék szerinti felismerése, mint az izolációban ejtett hangzóké. A különbségek statisztikailag ugyan nem igazolhatók, általánosságban azonban mégis úgy látszik, hogy nagyobb eltérés jelentkezett a „V” és „CVC”, illetve a „Vágott” és a „CVC” minták felismerési aránya között, mint a „V” és „Vágott” minták között.

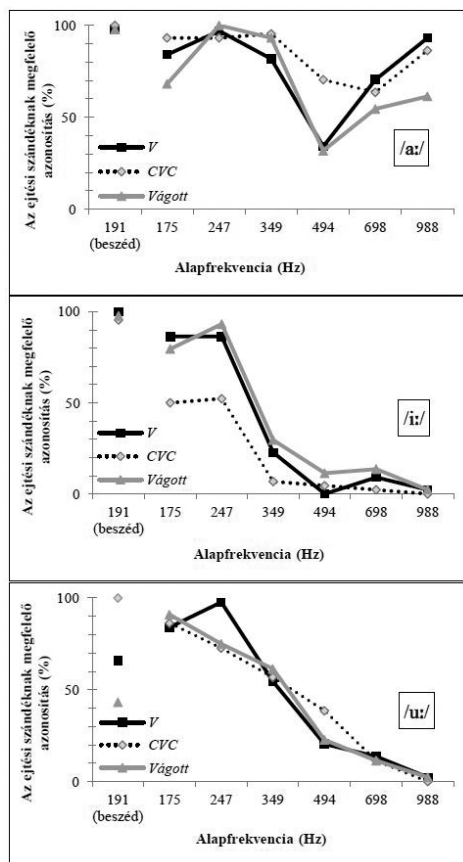
Bár az éneklésben nem láttuk igazolódni a fonetikai kontextus (a mássalhangzó-környezet és a zöngéindulás) hatását, a beszédben az elvárt tendenciákat találtuk: itt a legmagasabb, az ejtési szándék szerinti azonosítási arányt a „CVC” kondíció mutatta, melyet a „V” kondíció (izolált ejtés), végül pedig a „Vágott” kondíció (eliminált zöngékezdéssel előállított stimulusok) követték.

Ugyan a fonetikai kontextus hatása a magánhangzó-azonosításra nem igazolódott, az „ f_0 ”, a „kondíció” és a „magánhangzó-minőség” változókkal lefuttatott varianciaanalízis szerint nemcsak az f_0 , de a magánhangzó minősége [$F(2) = 22,673$; $p < 0,001$], illetve e két faktor kombinációja is [$F(2) = 7,419$; $p = 0,002$] befolyásolta a hangzóazonosítási értékeket. Ezért az ejtési szándék szerinti azonosítási arányok változását a magánhangzó-minőségek mentén is elemeztük.

A magánhangzó-minőségek közti összes lehetséges páros szembenállás összevetése szerint Bonferroni-korrekció után is szignifikáns különbség mutatkozott az /i:/ – /a:/ és /u:/ – /a:/ párok, tehát az első és negyedik nyíltsági fokú magánhangzók között ($p < 0,018$), de nem volt különbség a nyíltsági fok tekintetében egyező, első nyíltsági fokú (zárt, felső nyelvvállású) /i:/ és /u:/ között. Ezt az elkülönülést rajzolja ki az „ f_0 ” és az „ejtési szándék szerinti azonosítási arány” közti korreláció is, ha azt a magánhangzók csoportjain belül vizsgáljuk. A két változó ugyanis csak az /i:/ és /u:/ esetében mutat erős, szignifikáns összefüggést (/i:/: $r = -0,781$ $p < 0,001$; /u:/: $r = -0,900$; $p < 0,001$), míg az /a:/ esetében csak közepes, nem szignifikáns összefüggés látszik ($r = -0,384$; $p = 0,086$). A statisztikai elemzések, továbbá az adatokat vizualizáló 20. ábra tanúsága szerint tehát a zárt képzésű vokálisoknak az ejtési szándék szerinti azonosítása csökkent az alaphang emelésével, míg a nyílt képzésű (negyedik nyíltsági fokú) beszédhanghoz tartozó százalékok az f_0 értékétől megközelítőleg függetlenül alakultak.

Az /a:/-nak az ejtési szándék szerinti azonosítási aránya a *h'* zenei hang (494 Hz) alatt viszonylag magas volt minden kondícióban, a *h'* zenei hang magasságában azonban hirtelen esést tapasztaltunk (20. ábra). Ezen az alapfrekvencián 46%-ban azonosították a tesztelők a hangzókat az ejtési szándéktól eltérően, és minden esetben a magasabb első formánssal ejtett (zártabb képzésű) /b/-t jelölték meg válaszként. A *h'* fölött újabb emelkedés következett az ejtési szándék szerinti azonosítás arányában, majd ezt követően folyamatos emelkedést tapasztaltunk a legmagasabb alaphangig, ahol a „V” kondícióban (izolált ejtés) 93%, a „CVC” kondícióban (mássalhangzó-környezet) 86%, míg a „Vágott” kondícióban (eltávolított zöngé-

indítás) 61% volt az ejtési szándéknak megfelelő válaszok százalékos aránya. Az f'' zenei alaphang esetében ismét az / v / hang dominálta az ejtési szándéktól eltérő azonosításokat (az összes válasz 27%-ában), míg a legmagasabb, h'' zenei alaphang esetében az elől képzett / ε / (az összes válasz 8%-ában). Mindkét hangzó zártabb képzésű, és a beszédben magasabb első formánssal realizálódik, mint az ejtési szándékozott magánhangzó.



20. ábra: Az /a:/, /i:/ és /u:/ ejtési szándék szerinti azonosításának százalékos aránya az alapfrekvencia függvényében a három vizsgált kísérleti kondícióban

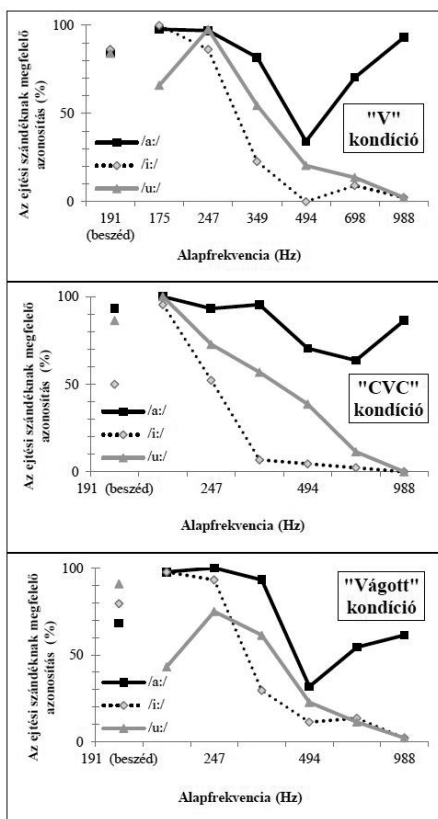
Az /i:/ az ejtési szándék szerinti azonosítása már az egészen alacsony alapfrekvenciájú, f' zenei hang magasságában hirtelen csökkenést mutatott (20. ábra); itt és a magasabb alaphangok esetében is legnagyobb részben / ε :-t jelöltek a hallgatók az ejtési szándéktól eltérő azonosítások alkalmával (f' : 80%, h' : 94%, f'' : 67% az összes válaszból). A h'' , azaz a legmagasabb zenei hang magasságában azonban az /i:/-t is leggyakrabban /a:/-ként ismerték fel (az összes válasz 80%-ában), az ejtési szándék szerinti azonosítások aránya pedig jóval a véletlen szintje alatt, 2% körül alakult (minden kondícióban).

Az /u:/ azonosítását leíró tendenciák a 19. ábrán bemutatott átlageredményekhez hasonlóan alakulnak: az f_0 emelésével fokozatosan csökkenő százalékarányokat láttunk, míg a „V” és „CVC” kondíciók között nem következetesen alakuló kapcsolatokat mutatkoztak. Az ejtési szándéktól eltérő azonosításokat tekintve a legtöbb alaphangfrekvencián a leggyakrabban előforduló hangzó az /o:/ (f és h : 13%, f' : 42%, h' : 62%, f'' : 25% az összes válaszból). Ugyanakkor a legmagasabb alaphangfrekvencián (h'') itt is az /a:/ uralta az ejtési szándéktól eltérő azonosítások hierarchiáját (az összes válasz 86%-ában), valamint az /u:/ azonosítása csak maximum 2%-ban (tehát jóval a véletlen szintje alatt) bizonyult az ejtési szándékkal megegyezőnek.

Az /u:/ beszélt ejtésben bemutatott realizációi (az /i:/-vel és /a:/-val szemben) a hipotézisben megfogalmazottaknak megfelelő tendenciát mutatták: a legnagyobb arányban a „CVC” kondícióbeli megvalósulást azonosították a hallgatók az ejtési szándék szerint, melyet a „V” és végül a „Vágott” kondíciók követtek. Azonban ez az /u:/ esetében is csak a beszélt ejtésre korlátozódó mintázat volt (a többi hang esetében pedig sem a beszédben, sem másutt nem tapasztaltuk ezt a tendenciát).

Összehasonlítottuk az összes lehetséges hangzószerkezetet a három kondíció alkotta csoporton belül is. Ezek az összehasonlítások megmutatták az egyetlen olyan tendenciát, mely a jelen adatokban a kondícióval és azon belül a mássalhangzó-környezet meglétével is összefüggésbe hozhatónak bizonyult. Egyik csoporton belül sem találtuk ugyanis eltérőnek a különböző nyíltsági fokú magánhangzókat, kivéve a „CVC” csoportot. Itt az eltérő nyíltsági fokú magánhangzók az ejtési szándéknak megfelelő azonosítási arányát eltérő mértékben látszott befolyásolni az f_0 értéke – bár csak a nyelv vízszintes mozgásának tekintetében jobban eltérő pár, az /i:/ és az /a:/ esetében találtunk a Bonferroni-korrektúra után is szignifikáns ($p < 0,02$) eltérést (a nyelv vízszintes mozgásának tekintetében egymáshoz közelebb álló /u:/ és /a:/ nem tért el). Az összehasonlításokat a 21. ábra grafikonjain szemléltettük.

Az adatokból jól látható, hogy a különbségek megjelenése a „CVC” kondícióban (illetve meg nem jelenése a többi kondícióban) annak köszönhető, hogy a „CVC” kondícióban tapasztalható leginkább az /a:/ percepciók stabilitása, ugyanakkor a zárt /i:/ is itt mutatja a legalacsonyabb, az ejtési szándék szerinti azonosítási arányokat. Míg a „V” és „Vágott” kondíciókban csak a h' zenei hang fölött, addig a „CVC” kondíciókban már a jóval alacsonyabb f hang fölött is divergáló tendenciákat látunk a három magánhangzóra.



21. ábra: A magánhangzók az ejtési szándék szerinti azonosításának aránya a három kondícióban

4.3. Következtetések

A jelen fejezetben bemutatott kísérletnek két célja volt: egyfelől a mássalhangzó-környezet, másfelől pedig a zöngéindítás mint az ejteni szándékozott magánhangzó-minőség felismerését segítő akusztikus kulcsok tesztelése. Amint azt a 4. fejezet bevezetőjében is hangsúlyoztuk, a mássalhangzó-környezet hatásának kérdése valójában kettős, és a jelen vizsgálat a két aspektus közül csak az elsőre fókuszált. Első kérdésünk ennek értelmében az volt, hogy valóban segíti-e, és ha igen, milyen mértékben a mássalhangzó-környezet a koartikulációs formánsátmenetek révén a magas alapfrekvencián énekelt magánhangzók az ejtési szándék szerinti azonosítását egy, a csak a bottom-up beszédészlelési folyamatokra, azaz az akusztikai elemzésre támaszkodó percepciósi feladatban. A vizsgálat második kérdéseként korábbi

pszichoakusztikai kísérletek eredményeire alapozva azt fogalmaztuk meg, hogy vajon az izoláltan ejtett magánhangzók zöngéindulási szakasza is hozzásegítheti-e a hallgatókat az ejteni szándékozott magánhangzók minőségének azonosításához. A vizsgálat elején azt feltételeztük, hogy mivel az alapfrekvencia emelésével egyre fokozódik a spektrális alul-mintavételezettség, valamint azért, mert a magas alapfrekvenciákon énekelt magánhangzók esetében $F_1 : f_0$ hangolást is feltételezünk, magas alapfrekvencián sem a formánsátmenetek, sem pedig a zöngéindulás nem segítheti a magánhangzók ejtési szándék szerinti azonosítását, hiszen az éneklésben magas alapfrekvencián az ejtési szándéktól eltérő akusztikai szerkezetű, illetve az f_0 emelésével egyre bizonytalanabb minőségű hangzók valósulnak meg. Az alacsonyabb alapfrekvenciák (azaz a beszédbeli megvalósuláshoz hasonlóbb akusztikai minőségű énekelt magánhangzók) esetében ugyanakkor azt is feltételeztük, hogy a mássalhangzó-környezet bizonyos mértékű pozitív hatása kimutatható lesz a közrezárt magánhangzónak az ejtési szándék szerinti azonosításában az izolált ejtésű magánhangzókhoz képest. A zöngéindítás pozitív hatását a mássalhangzó-környezet hatásánál valamelyest enyhébbnek feltételeztük, ugyanakkor a zöngéindítás nélkül prezentált hangok azonosításához képest a zöngéindítást is tartalmazó stimulusok esetében valamivel magasabb arányban vártuk a magánhangzók az ejtési szándék szerinti azonosítását az alacsonyabb alapfrekvenciákon.

A vizsgálat eredményei – SMITH és SCOTT (1980) hasonló vizsgálatával ellentétesen – igazolták azt a feltevésünket, mely szerint pusztán a mássalhangzó-környezetben rejlő dinamikus akusztikai kulcsok, azaz a formánsátmenetek nem elegendők ahhoz, hogy a hallgatók az ejtési szándék szerint azonosítsák az alapfrekvencia emelésével egyre jobban megváltozó, és így egyre bizonytalanabb minőségű énekelt magánhangzókat. A zöngéindítás a jelen adatok alapján (melyben csak a magánhangzótér szélső pontjait képező /a: i: u:/ magánhangzókat vizsgáltuk) csak az /u:/ beszédbeli megvalósulásának esetében tűnt az ejtési szándék szerinti azonosítást segítő akusztikai kulcsnak. Ráadásul eredményeink szerint mind a mássalhangzó-környezet, mind pedig a zöngéindítás hatása inkonzisztensnek látszott az izolált ejtésű magánhangzókvaló összevetésben. Az eltérések következtelensége azt jelenti, hogy nemcsak hogy adataink nem demonstrálták a CVC kontextusnak, illetve a zöngéindulásnak a pozitív hatását, hanem néhány esetben egyenesen annak az ellenkezőjét, azaz a kontextus és a zöngéindulás negatív befolyását is illusztrálták.

Az eredményeket a vizsgálat kiinduló feltételezésével összhangban a magas alaphangon ejtett (zöngés) hangok spektrális alul-mintavételezettségével, valamint az operaéneklésnek a magánhangzók képzését érintő sajátosságaival (azaz a magánhangzók az f_0 -tól függő ejtésével) magyarázhatjuk. Mivel az alapfrekvencia emelésével e két tényező következtében a magánhangzók akusztikai megvalósulása egyre inkább eltér a hangzók beszédbeli megvalósulásától, így a hangzók minősége egyre bizonytalanabb lesz, tehát nem is lehetséges az ejtési szándék szerinti azonosítás magas alapfrekvencián még az egyébként az azonosítást bizonyos más körülmények között segítő akusztikai kulcsok jelenléte esetén sem. Ugyanakkor már önmagában az is kérdés, hogy egyáltalán milyen mértékben segíthetnének a formánsátmenetek akkor, ha a magánhangzók akusztikai szerkezete nem is változna, hiszen

a spektrális alul-mintavételezettség nemcsak a vokálisok tiszta fázisát jellemzi, hanem minden olyan beszédhangot és hangrészletet is, melynek akusztikai lenyomatát a hangszalagműködés következtében létrejövő zöngé hozza létre. Ennek megfelelően ugyanis a zöngés mássalhangzók és a bármely két szonoráns között realizálódó formánsátmenetek is alul-mintavételezetté, azaz fonetikailag alulspecifikálttá válnak. A mássalhangzókkal kapcsolatban ráadásul még arról az eddig alig kutatott kérdéstről sem szabad megfeledkeznünk, hogy vajon mi jellemző ezekre a beszédhangokra az éneklésben. Vajon az éneklés mint speciális beszédképzési mód hatással van az énekelt mássalhangzók artikulációs és akusztikai megvalósítására is?

Az operanéklés (vagy másként a *bel canto* technika) egyik alapvetése a magánhangzók időtartamának maximalizálása a mássalhangzók időtartamának rovására (lásd pl. Váradi 2010), amely megszorítás azt jelenti, hogy bár a technika elsősorban a magánhangzók ejtésére fókuszál, közvetett módon a mássalhangzók ejtését is meghatározza: azok redukcióját írja elő. Ennek a mássalhangzó-redukciónak a módját és jellegét azonban, a jelen értekezés szerzőjének ismeretei szerint, kísérletes úton eddig még csak egyetlen alkalommal vizsgálták. Ebben az elemzésben DEME, GRÁCSI és JANKOVICS (2013) az énekelt obstruensek zöngésségének alakulására fókuszált. A szerzők egyfelől bemutatták az egyik vizsgált zöngés explozíva (a /b/) rövidülését szókezdő és szózáró helyzetben, másfelől pedig leírták több szókezdő, szó belseji és szózáró explozíva és frikatíva zöngésségének megváltozását is. Ezen túlmenően azonban a szerzők arról is beszámoltak, hogy a mássalhangzók címkézésekor gyakori problémát jelentett a vizsgált explozívak felpattanásainak elmaradása (vagy legalábbis detektálhatatlansága a spektrogramon és az oscillogramon), illetve a réshangok turbulens zörejjössze tevőinek alacsony intenzitása.

Az obstruensek lenizációjának, illetve az ebből következően alacsony intenzitású (sőt néhol egyenesen hiányzó) akusztikai összetevőknek az észlelés szempontjából két okból is jelentősége lehet. Egyrészt valószínűsíthető, hogy az ily módon megváltozott akusztikai szerkezetű mássalhangzók azonosítása is nehezebb lesz az ejtési szándék szerint. Másrészt pedig az is feltételezhető, hogy a megváltozott módon ejtett mássalhangzók artikulációja a koartikuláció révén a magánhangzók ejtésére is hatással van, amivel együtt természetesen a koartikuláció akusztikai lenyomatát, azaz a koartikulációs formánsátmenetek megvalósulását is befolyásolja. Ha ugyanis a magánhangzót megelőző explozíva felpattanása, illetve a frikatíva zörejes eleme elmarad, a formánsátmenetek sem olyan módon jelennek meg (vagy nem is jelennek meg egyáltalán) a magánhangzó akusztikumában (és az azt vizualizáló spektrogramban), mint egy gondosan artikulált explozíva/frikatíva esetén, hiszen a nyelv feltehetőleg csak részben teszi meg (vagy esetleg egyáltalán nem) a felpattanás/réses zörej létrehozásához egyébként bejárat utvonalat a szájüregben. Mindez pedig egyúttal azt is jelenti, hogy az éneklésben a nem gondosan artikulált obstruensek környezetében elhelyezkedő magánhangzóknak és a mássalhangzó-magánhangzó átmenetekben a beszédben ejtett beszédhangoknál kevesebb dinamikus információ rejlik, tehát kevesebb a percepciót elméletben támogatni képes „akusztikus mankó” is. Okfejtésünkben egyértelműen következik, hogy

a koartikulációs formánsátmenetek hatásának vizsgálatához az énekelt mássalhangzók realizációjának és esetleges redukciójának kutatása is szükséges. Addig ugyanis nincs okunk feltételezni, hogy a mássalhangzó, illetve a koartikulációs formánsátmenetek bármilyen befolyással bírhatnak a percepciós működésekre az énekelt magánhangzók esetében, amíg az sem tisztázott, hogy egyáltalán maga a konzonáns és a koartikuláció, valamint a koartikulációs formánsátmenetek megjelennek-e, és így elérhetőek lesznek-e az akusztikumban, illetve a percepció számára. Az éneklésben realizálódó mássalhangzók akusztikai és lehetőség szerinti artikulációs vizsgálata a jövőbeni terveink között szerepel.

Az itt bemutatott kísérlet eredményeihez kapcsolódva fontos kiemelni azt is, hogy az adataink kimerítő értelmezéséhez a jövőben szükséges lesz a különböző képzéshelyű mássalhangzók (tehát nem csak a bilabiális, de az alveoláris és veláris felpattanók) vizsgálata is, melyek képzéshelyük (illetve eltérő frekvenciaszerkezetük) következtében eltérő formánsátmenetekkel realizálódnak a CVC szekvenciákban. A mássalhangzó képzéshelye és a közrezárt magas alapfrekvenciájú énekelt magánhangzók azonosítása közti összefüggések vizsgálata szintén a jövőben megvalósítandó terveink közé tartozik.

A magánhangzó-azonosítási arányok kondíciók szerinti összehasonlításakor találtuk meg a mássalhangzó-környezettel összefüggésbe hozható egyetlen percepciós mintázatot: a magánhangzók ugyanis a CVC kondícióban mutatták a legszéttartóbb azonosítási tendenciákat. Az /a:/ esetében a CVC kondícióban tapasztaltuk a legmagasabb, míg az /i:/ esetében a legalacsonyabb arányokat; a két, a nyíltsági fok mentén eltérő hangzó észlelési tendenciái ebben a kondícióban tértek el a legnagyobb mértékben. Bár e magánhangzók frekvenciaszerkezetbeli eltérései miatt a magánhangzók észlelésében is különbségeket várunk, az itt említett eredmény jelentőségére pusztán a jelen vizsgálat adatai alapján nem következtethetünk.

Kiemelendő, hogy a produkció módja (az ének mód, illetve a beszéd mód közti váltás) két okból is az észlelést befolyásoló tényezőnek tűnik. Egyfelől az énekelt mintákkal szemben a beszédben ejtett hangzónál több esetben is (az összesített eredményekben, illetve az /u:/ esetében is) láttuk azt a korábbi szakirodalom által feltételezett sajátosságot, mely szerint a mássalhangzó-környezet és a zöngéindulás pozitív hatást fejt ki a magánhangzónak az ejtési szándék szerinti azonosítására a beszédben. Másfelől pedig az /a:/ és /i:/ esetében a beszélve ejtett realizációk mutatták a legmagasabb azonosítási arányokat (alaphangmagasságtól függetlenül) még azokhoz az énekelt alaphangmagasságokhoz képest is, melyek a beszéd átlagos f_0 -értékéhez közeliak voltak.

Amint az összegzés elején már megemlítettük, a jelen kutatás eredményei ellentmondanak a mássalhangzó-környezetet vizsgáló egyetlen korábbi tanulmányban (SMITH–SCOTT 1980) olvashatóknak. Megválaszolendő kérdés, hogy mivel magyarázható az ellentmondás. Amint arra már több ponton is utaltunk értekezésünkben, SMITH és SCOTT (1980) eredményei az adatokat befolyásoló változók nem szigorú kontrollálása következtében nehezen (vagy talán egyáltalán nem) interpretálhatók. A szerzők ugyanis az általuk vizsgált két kondícióval két jelentősen eltérő percepciós működést vetettek össze, melyek ráadásul jelentősen eltérő számú válaszlehetőséget engednek meg az azonosítási feladat során. Míg a mássalhangzó-környe-

zetben a kísérleti személyek értelmes szavakat kellett, hogy azonosítsanak, ami gyakorlatilag négy lehetőség közti választást jelentett, addig az izolált ejtésben magánhangzókat kellett felismerniük, tehát az utóbbi esetben mindösszesen az elemi hallási észlelés és az alsóbb szintű (akusztikai elemzésre hagyatkozó) beszédészlelési folyamatokra támaszkodhattak a válaszadáskor, mely során így az angol nyelv összes szóba jöhető monoftongus magánhangzója (azaz legalább 10 lehetőség) közül kellett választaniuk. SMITH és SCOTT (1980) vizsgálatában ily módon vegyült a mássalhangzó-környezet és a magánhangzó-azonosítás problémájának két, egymástól elkülönítendő, hiszen alapvetően eltérő aspektusa, ezért a vizsgálat eredményei valójában sem az egyik – a formánsátmenetek mint akusztikus kulcsok hatásával kapcsolatos –, sem pedig a másik – a magasabb szintű percepciós működések aktivációjával kapcsolatos – problémakörre nem engednek következtetéseket levonni. Az ebben a fejezetben bemutatott vizsgálat ezzel szemben az itt említettek közül kizárólag az első kérdésre fókuszált (a kérdés megválaszolásához jól kontrollált feltételek mellett kapott eredményeket szolgáltatva). Ebben a vizsgálatban azt a korábbi észlelési tesztlejzerekkel is összehasonlítható következtetést vonhattuk le, hogy a mássalhangzó-környezet nem segíti a magas alaphangfrekvencián énekelt magánhangzók számára az ejtési szándék szerinti azonosítását a koartikulációs formánsátmenetek révén.

A következő fejezetben bemutatásra kerülő kísérletünkben – bizonyos értelemben tehát a fent ismertetett kísérlet folytatásaképp – a mássalhangzó-környezet hatásának másik aspektusát, a magánhangzók értelmes szavakba ágyazottságának hatását teszteljük és vetjük össze a mássalhangzó-környezet mint dinamikus akusztikai kulcsok társaságának hatásával az énekelt magánhangzók észlelésében. Reményeink szerint e két kérdés szétválasztása és összehasonlítása megerősíti azt a javaslatunkat, mely szerint a kontextus hatásának kérdése az észlelésben valójában két, egymástól élesen elkülönítendő fogalmat jelöl, és magyarázatot szolgáltat arra is, hogy miért kaptunk a jelen kísérletben SMITH és SCOTT (1980) eredményeitől jelentősen eltérő adatokat.

5. A mássalhangzó-környezet hatása az énekelt magánhangzók észlelésére: a magánhangzók azonosítása értelmes és értelmetlen szavakban énekes és nem énekes hallgatók percepciójában

A magas alapfrekvenciákon énekelt magánhangzókat sajátos artikulációs és akusztikai jellemzőik miatt bizonyos értelemben fonetikailag alulspecifikáltnak tekinthetjük. Az alulspecifikáltság két aspektusa a) az artikulációs és akusztikai célkonfiguráció megváltozása az $F_1 : f_0$ formánshangolás megjelenésének következtében (azaz az F_1 formáns az éneklés alapfrekvenciájára hangolása akkor, ha a megemelt f_0 egyébként magasabb lenne, mint a magánhangzónak a beszédben jellemző F_1 formánsa), valamint b) a magánhangzó-realizációk spektrumának alul-mintavételezettsége (azaz a spektrális tér kitöltöttségének csökkenése). Mind a célkonfiguráció megváltozásának mértéke, mind pedig a spektrális alul-mintavételezettség fokozódik az alapfrekvencia emelésével.

A fonetikai alulspecifikáltság miatt kérdés, hogy miként és milyen alapfrekvenciáig ismerhetők fel az énekelt magánhangzók az ejtési szándék szerint; valamint az is, hogy létezhetnek-e olyan „percepciósmankók”, melyek az ejtési szándék szerinti azonosítást támogatják. Amint azt már az értekezés korábbi pontjain bemutattuk, a beszédpercepció szakirodalmára támaszkodva az énekelt magánhangzók vizsgálatában sem ritka, hogy egyes kutatók a magánhangzók azonosítását segítő körülményként a mássalhangzó-környezet hatását feltételezzék. Ugyanakkor azt is láttuk, hogy e feltételezés az éneklésre vonatkoztatása valójában nélkülözi az empirikus megalapozottságot, hiszen az egyetlen, a kérdést vizsgáló kísérlet eredményei a kísérlet módszertani megoldásai (és a *confound* hatás) miatt csak korlátozottan értelmezhetők. A mássalhangzó-környezet hatásának kérdése az énekelt magánhangzók azonosításában tehát nem tekinthető egyértelműen tisztázottnak.

Az előzőekben már ismertettük azt a feltételezésünket, mely szerint a mássalhangzó-környezet hatását vizsgáló vagy feltételező szakirodalom problémájának egyik fő forrása a mássalhangzó-környezet, illetve a kontextus fogalmának pontatlan meghatározása, használata. Azt tapasztalhatjuk ugyanis, hogy a feltételezések alapvetően a mássalhangzók és a kérdéses magánhangzó közti formánstemenetek mint akusztikus kulcsok szerepére hivatkoznak a magánhangzó-azonosításban, miközben az egyetlen korábbi vizsgálat inkább csak a kontextus mint értelmes szavakba ágyazottság hatásával kapcsolatban enged következtetéseket megfogalmazni. SMITH és SCOTT (1980) kísérlete, bár a szerzők szándéka szerint a koartikulációs formánstemenetek, illetve a dinamikus akusztikai kulcsok vizsgálatát célozza meg, valójában egészen más kérdést vizsgál. Eredményeik ugyanis nem a dinamikus kulcsok, hanem

a magánhangzók értelmes szavakba ágyazottságának pozitív hatását igazolják, ráadásul kizárólag egy olyan elrendezésben, melyben a válaszlehetőségekként felkínált értelmes szavak száma jóval kisebb, mint az összehasonlítás alapját képező másik kondícióban rendelkezésre álló válaszok (azaz az izolált ejtésű magánhangzók vagy „értelmetlen szavak”) száma.

Éppen ezért a jelen kötetben a két kérdés szétválasztott vizsgálatát, valamint a kétféle hatás és a kétféle percepció működés összevetését végezzük el: az előző fejezetben a dinamikus akusztikus kulcsok, ebben a fejezetben pedig az értelmes szavakba ágyazottság hatását teszteltük SMITH és SCOTT (1980) eredményeiből kiindulva. SMITH és SCOTT (1980) vizsgálata bizonyította, hogy a magánhangzók értelmes szavakba ágyazottan magasabb arányban ismerhetők fel az ejtési szándék szerint akkor, ha a szóba jöhető értelmes szavak, azaz a lehetséges magánhangzóválaszok száma kisebb, mint amennyit az értelmetlen szavak és/vagy izolált magánhangzók azonosításának feladata megenged – ez utóbbit ugyanis csak a magánhangzó-készlet mérete befolyásolja. Az előző fejezetben ismertetett kísérletünk szerint ugyanakkor a mássalhangzó-kontextus pusztán a dinamikus akusztikai kulcsok révén nem segíti a magánhangzók az ejtési szándék szerinti azonosítását a magas alapfrekvenciájú énekülésben (ahol csak az akusztikus elemzést végző bottom-up folyamatokra támaszkodhat a percepció).

Míndezen alapján már csak az a kérdés vár megválaszolásra, hogy miként hat a magánhangzók felismerésére az, ha a felismerési feladatban olyan módon hallják azokat értelmes szavakba ágyazottan az adatközlők, hogy a szóba jöhető értelmes szavak száma nem jelentősen kisebb (közel ugyanolyan nagyszámú), mint az értelmetlen szavakban szóba jöhetőké. Az itt bemutatásra kerülő vizsgálatban ezt a kérdést járjuk körül.²⁴

A jelen fejezet tehát egyfelől arra a kérdésre keresi a választ, hogy különbözik-e egymástól a mássalhangzó-környezetnek mint a percepciót segítő akusztikai kulcsok tárának és a mássalhangzó-környezetnek mint a magasabb észlelési folyamatokat és a nyelvi tudást is aktiváló körülménynek a hatása. A vizsgálatban emellett azonban egy másik, az énekelt magánhangzók azonosításával kapcsolatban gyakran felmerülő kérdést is tesztelni kívánunk. Percepció kísérleteink során gyakran találkozunk énekesekkel és énekelni nem tudó adatközlőkkel is, akikkel a kutatás kapcsán az éneklésről, a beszédképzés folyamatairól és a beszédészlelésről is beszélgetünk. Az ilyen alkalmak során számos, a beszéd- és énekhangképzést érintő naiv elképzelést ismerünk meg. Ezek közül egy (viszonylag gyakran előkerülő) intuitív feltételezés szerint az énekesek a nem énekeseknél a percepció tekintetében is jobban teljesítenek, azaz – e feltételezések szerint – a nem énekesekkel szemben az énekesek nemcsak beszédet képezni (azaz beszédhangokat ejteni) képesek magas alapfrekvencián, de ugyanezen beszédhangoknak az ejtési szándéknak megfelelő azonosításában is jobban teljesítenek.

A kísérleti résztvevők általában arra alapozzák ezt a feltevésüket, hogy úgy vélik, a nem énekesekkel ellentétben az énekesek pontosan tudják, hogy milyen artikulációs-akusztikai módosításokat hajtanak végre egy-egy magánhangzó ejtésekor a magas és szép (azaz homo-

²⁴ Köszöneti illeti Máty Katalint, aki egyfelől megerősített abban, hogy a kutatási kérdés releváns problematikát feszeget (mely még a mindennapi beszéd észlelésének vonatkozásában sem tekinthető triviálisnak), másrészt javaslataival és kritikai észrevételeivel segített a kísérlethez használt nyelvi anyag kidolgozásában is.

gén hangszínezetű) énekehang létrehozásának érdekében, éppen ezért arra is képesek, hogy a módosításokról szóló tudásuk alapján visszakövetkeztessenek az eredetileg ejteni szándékolt hangra. Érdekes módon ez a talán naivnak ható elképzelés jól egybeesik a beszédpercepció szakirodalmából ismeretes motoros elmélettel (vö. a jelen értekezés *1.2.2.1. alfejezetét*). Emlékeztetőül, a motoros elmélet központi feltételezése az, hogy a beszédészlelés alapvetően nem az akusztikai mintázatok felismerésén, hanem (az akusztikai jelekből kikövetkeztethető) artikulációs gesztusok azonosításán múlik. Az elmélet követői úgy vélik ugyanis, hogy az agy nem a hangok akusztikai lenyomatát, hanem a beszédhangok ejtéséhez szükséges motoros parancsokat (artikulációs gesztusokat) reprezentálja (mégpedig invariáns módon), és az észlelt akusztikai jel is, gyakorlatilag automatikusan, ilyen információegységekké transzformálódik az észlelési folyamatok során (LIBERMAN–MATTINGLY 1985). A motoros elméletből következően a megvalósítás stílustól, beszédhelyzettől függően változatos lehet, de mégis minden megvalósításban benne rejlenek a változatlan alapegységek, mégpedig a produkcióhoz szükséges artikulációs gesztusok, amelyek az észlelést is lehetővé teszik.

A motoros elmélet tehát bizonyos tekintetben megerősíti azokat az intuitív vagy laikusnak tekinthető feltételezéseket, mely szerint az énekesek saját artikulációs gyakorlatukból kifolyólag a nem énekes hallgatónál sokkal inkább képesek a jelentős akusztikai aluspecifikáltsággal megjelenő hangzók ejtési szándék szerinti azonosítására. A motoros elméletet sokan sok okból vitatják,²⁵ és a jelen értekezésben nem áll szándékunkban állást foglalni e vitában. Mégis, ha azt a feltételezést nem is fogadjuk el, hogy az artikulációs gyakorlat közvetlenül összeköttetésben volna a percepció mechanizmusaival, az továbbra is kérdés marad, hogy az énekesek artikulációs tudatossága működésbe léphet-e, és ha igen, pozitívan járul-e hozzá az énekeseknek a magas alaphangfrekvencián énekelt magánhangzók azonosításában nyújtott teljesítményéhez. A vizsgálat második kérdése tehát az, hogy vajon jobban teljesítenek-e az énekesek az énekelt magánhangzóknak az ejtési szándék szerinti azonosításában, mint az éneklésben nem gyakorlott észlelők. Ismereteink szerint ezt a kérdést korábban még nem vizsgálta tudományos kísérlet.

A fejezetben bemutatásra kerülő kísérletek célja tehát kettős. Az első cél az eltérő percepció működéseket feltételező feladattípusokban (azaz az értelmes és értelmetlen szavakban) tapasztalható magánhangzó-azonosítási tendenciák összehasonlítása, ezzel is tovább árnyalva a mássalhangzó-kontextus hatásának kérdését a magas alaphangfrekvencián énekelt magánhangzók észlelésében. Eddigi akusztikus és észlelési vizsgálataink alapján azt feltételezzük, hogy nem hatékonyabb az ejtési szándéknak megfelelő magánhangzó-azonosítás az értelmes szavakban akkor, ha az értelmes és értelmetlen szavakban ejtett magánhangzók azonosítását

²⁵ Egyrészt a motoros elmélettel szemben állnak az akusztikum elsődlegességét hangsúlyozó, illetve az agyi reprezentációk variabilitását feltételező fonetikai teóriák (utóbbira példa LINDBLOM 1990-ben közölt H&H elmélete, vö. még a jelen dolgozat *1.2.2.2. alfejezete*). Másrészt pedig idegtudományi vizsgálatok is cáfolják a motoros teóriát: HICKOK (2010) például az agyi aktivitás és a tükkörneuronok vizsgálata alapján azt feltételezi, hogy az észlelés közben tapasztalható motoros agyi aktivitás valójában a szenzoros területekről terjedő aktiváció eredménye, amelynek célja nem az észlelési folyamatok segítése, hanem a megszólalásra, a beszédprodukcióna való felkészülés.

nem jelentősen eltérő számú (de akár nem is szükségszerűen ugyanannyi) válaszlehetőséget megengedő kondíciókban vetjük össze. Feltevésünk szerint ugyanis a kontextus pozitív hatása nem a koartikulációs formánsátmenetekben, hanem az értelmes szavak által korlátozott válaszlehetőségekben rejlik, mely hatás közel kiegyenlített számú válaszlehetőségek esetén erősen lecsökken (vagy teljesen meg is szűnik). A kísérletünk másik célja az énekes és nem énekes hallgatók teljesítményének összehasonlítása a magánhangzó-észlelés tekintetében. Elfogadva a motoros elmélet egy lehetséges értelmezéséből is kiinduló feltételezést, mely szerint az énekeseknek a saját artikulációs működéseikkel kapcsolatos tapasztalata vagy tudása a beszédpercepció valamely szintjén képes beépülni az észlelési folyamatokba, és így percepció előnyt jelent, azt feltételezzük, hogy az énekesek nagyobb arányban képesek az ejtési szándéknak megfelelő hangzóazonosításra, mint a nem énekes hallgatók.

5.1. Anyag, módszer és kísérleti személyek

A hipotézisek vizsgálatára egy két ülésből álló percepció tesztet terveztünk, melyet 20 „énekes” (26 és 45 év közöttiek, átlagos életkoruk 31 év) és 20 „nem énekes” (26 és 52 év közöttiek, átlagos életkoruk 29 év) adatközlővel vettünk fel. Mindannyian egészséges, ép hallású, magyar anyanyelvű beszélők voltak. Az énekesek csoportjában 15 profi szoprán (nő), 2 profi tenor és 3 profi bariton (férfi) képviselte a magas alaphangfrekvenciákhoz szokott, klasszikus nyugati típusú hangképzést tanult operaénekeseket. A „nem énekes” csoportban fonetikailag, valamint beszéd- és énektechnikailag képzetlen, „naiv” adatközlők szerepeltek.

A percepció kísérlethez két stimuluskészletet állítottunk elő, majd a stimulusokat egy 29 éves profi szoprán énekesnő ejtésében rögzítettük csendesített szobában, gömbkarakterisztikájú mikrofon segítségével. Az egyik stimuluskészlet az értelmes szavakat tartalmazta („Szavak” kondíció), míg a másik az értelmetlen szavakat („Álszók” kondíció). Mindkét stimuluscsoport CVC szerkezetű szekvenciákból állt, melyekben a magánhangzó-készlet maximális kihasználása, valamint a koartikulációs hangátmenetek egységesítése miatt az /b a: e i: ø: y:/ magánhangzók szerepeltek alveoláris képzéshelyű mássalhangzók között. A „Szavak” kondícióban a vizsgált magánhangzókat a *szVr* mássalhangzóvázba ágyaztuk (pl. *szár*), az „Álszók” kondícióban pedig a *dVr* kontextust használtunk (pl. *dár*). A stimulusokat az énekes hat, egymástól (a temperált hangolás szerint) egyenlő távolságra lévő alaphangfrekvencián énekelte fel kitarva (a hangsor teljes időtartama alatt állandó alaphangfrekvencián), a számára kényelmes hangerősséggel (és vibratóval), 175 Hz és 988 Hz között (*az f, h, f', h', f'', h''* zenei hangokon, lásd 6. táblázat), ezzel gyakorlatilag teljesen lefedve a szoprán hangosztály alaphangtartományát. Az énekelt stimulusok mellett a hangsorokat az énekes beszélt ejtésében is rögzítettük. Így összesen ($7 f_0 \times 2$ kondíció $\times 6$ magánhangzó) 84 teszt-hangsort vizsgáltunk.

A percepcióos tesztet egy, a Praat programban (BOERSMA–WEENINK 2009) írt és futtatott szkripttel vezéreltük. A teszthangsorokat a kondíciók szerinti bontásban összesen 42 disztraktor szóval és álszóval együtt, két részben mutattuk be a résztvevőknek. Minden résztvevő minden hangsort kétszer hallott, de minden résztvevő más-más (randomizált) sorrendben hallotta a teszthangsorokat, így az esetleges fáradás következtében megjelenő hatások eloszlottak a stimulusok között. Az első részben minden esetben a „Szavak” kondíció stimulusait tartalmazó tesztet végeztük el az adatközlőkkel, majd pár perc szünet beiktatása után ezt az „Álszók” kondíció elemeit tartalmazó rész követte.

A résztvevők feladata mindkét tesztfázisban ugyanaz volt: ki kellett választaniuk a képernyőn ortografikus alakban megjelenő kilenc magyar magánhangzó közül azt, amelyiket hallani vélték. A válaszlehetőségek a hosszú (vagy fonémakategória-váltás nélkül hosszan ejtendő) magyar magánhangzók voltak: /ɒ a: e e: i: o: ø: u: y:/. Az adatközlőknek szóló instrukciók azonban némileg eltértek a két tesztben. Míg az első részben felhívtuk az adatközlők figyelmét arra, hogy értelmes magyar szavakat fognak hallani, és ezeket kell felismerniük, addig a második részben nyomatékosan megkértük őket, hogy kifejezetten az elhangzó magánhangzó minőségére figyeljenek, ugyanis értelmes és értelmetlen hangsorok egyaránt előfordulhatnak a bemutatott stimulusok között. Az egyes hangsorok elhangzásakor a szkript a képernyő felső sávjában mindkét feladatban megjelenítette az azonosítandó magánhangzót közrefogó mássalhangzó-kontextust, így maximalizálva a koartikulációs hangátmenetektől eredhető pozitív hatást. Ezen túlmenően pedig az értelmes szavakat tartalmazó feladatban ugyanitt a feladatmegoldás során elfogadott, legitim válasznak számító értelmes szavakat is kivetítettük. Ezzel a feladat az instrukcióknak megfelelő elvégzését biztosítottuk, hiszen az elfogadott válaszok vizuális bemutatásával minden bemutatás alkalmával emlékeztettük az adatközlőket a feladatban legitimnek tekintett válaszokra.

Tekintve, hogy az /o: u:/ magánhangzók az „Álszók” *dVr* mássalhangzózázában is értelmes szót alkotnának, ezért ezeket nem vizsgáltuk, sőt egyúttal ki is zártuk a kísérletből úgy, hogy az első feladatban ezeket illegitim válaszoknak minősítettük. Ehhez az /o:/-t és /u:/-t tartalmazó értelmes szavakat (*szór, szúr*) nem soroltuk fel a lehetséges (elfogadható) válaszok között a képernyő felső részében az első feladatblokkban. Mivel pedig az /e:/ a „Szavak” *dVr* mássalhangzózázában is értelmes szót alkotott volna, így ebben a kísérletben ezt a beszédhangot sem vizsgáltuk. Ennek ellenére, amint azt fentebb említettük, válaszlehetőségként mindkét feladatban mind a kilenc különböző minőségű magyar köznyelvi magánhangzót kivetítettük, tehát egyfelől mindkét fázisban megjelenítettünk három olyan magánhangzót (az /e:/-t, /o:/-t és /u:/-t), mely (legalábbis elméletben) nem hangzott el, másfelől pedig ezzel az első fázisban olyan lehetőségeket is megjelenítettünk, melyek az ott közölt instrukciók szerint nem tartoztak a legitim megoldások közé. Emellett a kísérleti elrendezés mellett azért döntöttünk, mert fontosnak tartottuk fenntartani annak a lehetőségét, hogy a résztvevők a lehető leghababban „hibázhassanak” a feladatok végrehajtása során. Más vizsgálatokban kapott eredményeink alapján ugyanis például feltételezhető, hogy az /i:/ már igen alacsony alapfrekvencián is elveszíti eredeti perceptuális minőségét, és azt a hallgatók legnagyobb

számban /e:/ fonémaként kategorizálják (lásd 2. és 6. fejezet). Érdekesnek tartottuk tehát megfigyelni, hogy mi történik akkor, ha a feladatutasítás (mely szerint az /e:/ nem legitim válasz) és a hallott magánhangzó perceptuális fonémakategóriája (mely feltételezhetően már alacsony alapfrekvencián is az /e:/ hang megjelölésére sarkallná a résztvevőket) konfliktusba kerül.

A kísérlet itt kifejtett sajátosságának következtében azonban az eredmények taglalásakor mindig szem előtt kell tartanunk két fontos szempontot. Az egyik szempont az, hogy az egyes magánhangzókra kapott azonosítási arányokat ne abszolút módon értelmezzük, hanem annak is a szem előtt tartásával, hogy a „Szavak” kondícióban egy-egy magánhangzéválasz csekély mértékben, de nagyobb valószínűséggel jelenhetett meg. Ha a résztvevő a fent említett három magánhangzót a feladatutasításnak megfelelően kizárta a válaszlehetőségek közül, akkor ez a valószínűség a „Szavak” kondícióban 17%, míg az „Álszók” kondícióban 11% volt. Ha az adatközlő a legitim válaszokat jelző szólista figyelmen kívül hagyásával döntött, tehát az /o: u:/ magánhangzókat is legitim válasznak tekintette, a valószínűségek rendre 13% és 11% voltak. Ha pedig az adatközlő a „szabályok” teljes figyelmen kívül hagyásával végezte el a tesztet, a valószínűségek egyenlők (11-11%) voltak. A másik szempont mindennek megfelelően pedig az, hogy nem csak az ejtési szándéknak megfelelő válaszokra és a „helyes tippelésre”, de az ejtési szándéktól eltérő azonosítások bizonyos tendenciáinak megjelenésére is nagyobb esély lehetett a „Szavak” kondícióban. Ezzel a megoldással a kísérletben tehát valójában tovább megyünk annál, mintsem hogy csak az értelmes és értelmetlen szavakban megmutatózó magánhangzó-azonosítás jellemzőit vetnénk össze (erre elegendő lenne, ha 6-6 válaszlehetőséget hagynánk a résztvevőknek). Demonstrálni kívánjuk ugyanis azt is, hogy a két vizsgált feladat közti alapvető különbség is kizárólag a válaszlehetőségek számában rejlik. Eszerint (a valószínűségek természetéből fakadóan) nagy különbségeket okoz, ha kisszámú válaszlehetőség áll szemben nagyszámú válaszlehetőséggel (ti. a magánhangzórendszer összes magánhangzójával), ahogy SMITH és SCOTT (1980) kutatásában, de jóval kisebb különbséget, illetve a különbségek tökéletes eltűnését eredményezi már az is, ha az összevetett csoportokban hasonlóan nagy (de nem feltétlenül azonos) számú válaszlehetőség van.

A teszteket csendes körülmények között vettük fel, a hanganyagokat zárt fejhallgatón keresztül hallották a résztvevők.

Az eredményeket alapfrekvenciánként az összes magánhangzóra összesítve és az egyes magánhangzókra külön-külön is elemeztük a résztvevők két csoportjában, valamint megvizsgáltuk a nem az ejtési szándék szerint azonosított hangzók esetében fellelhető tendenciák típusait is a két kondícióban.

Adatainkon az R programmal (R Core Team 2013) végeztünk statisztikai elemzéseket: előbb Shapiro–Wilk-próbával ellenőriztük a normális eloszlás kritériumát, majd normális eloszlás esetén parametrikus teszteket (*t*-próba, párosított *t*-próba, ANOVA, ismételt méréses ANOVA), nem normális eloszlás esetén pedig nemparametrikus tesztet (Mann–Whitney *U*-próba) végeztünk.

5.2. Eredmények

5.2.1. Bevezető gondolatok a résztvevők benyomásai alapján

A teszt felvétele után minden adatközlőt arra kértünk, foglalja össze, milyen nehézségűnek találta a két tesztfeladatot, és hogy milyen általános benyomása született a saját teljesítményéről. A résztvevők mindkét csoportban egyöntetűen állították, hogy a két feladat között nem érzékeltek nehézségi különbséget, valamint hogy a magasabb alapfrekvenciákon ejtett hangzók esetében is jól teljesítettek, bár érzékelték, hogy „a magas hangoknál minden olyan á-szerű”. Az énekes csoport tagjai – a kísérletvezetőt informálandó – beszámolójukat a következőkkel is kiegészítették. Létezik egy, a kísérletben részt vevők tapasztalata szerint az elmúlt évtizedekben a magánének-tanításban gyakorolta alkalmazott, de mára már sokak által túlhaladottnak tekintett produkciós megoldás, amelynek értelmében az énektanárok a magas alapfrekvencián nehezen ejthető /i:/ hang helyett /y:/, míg az /e:/ helyett /ø:/ ejtésére „képezik ki” diákjaikat. A „csalás” célja, hogy az ajakkerekítéses ejtés révén az énekes számára könnyebb legyen a magas alaphangok elérése,²⁶ maga a produkciós stratégia pedig láthatólag közismert énekesi berkekben.

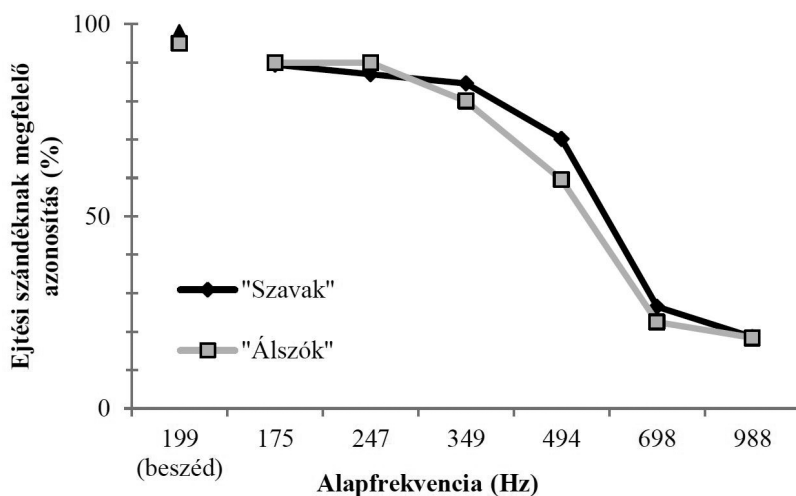
A válaszok fényében két fontos, a kísérlet eredményeit érintő megfigyelést tehetünk. Egyrészt fontos látnunk, hogy bármilyen artikulációs-akusztikai változáson esnek át a beszédhangok az éneklés során az alapfrekvencia függvényében, a résztvevők minden, a kísérletben elhangzott hangsorban létező magyar magánhangzókat véltek felfedezni, melyek számukra egyértelműen besorolhatónak tűntek egy-egy fonémakategóriába. Másrészt pedig az énekesek kommentárjai arra is felhívják a figyelmet, hogy érdemes lehet megfigyelni, milyen észlelési tendenciákat találunk a „nem énekes” és „énekes” résztvevőcsoportokban (és egymással összevetve) az /i: y: e: ø:/ magánhangzókkal kapcsolatosan.

²⁶ Ennek a stratégiának az oka feltehetőleg abban rejlik, hogy az énekesek igyekeznek kerülni a magasabb gégehelyezettel járó zöngékepzést, mert az könnyen vezethet feszített, préselt fonációhoz (ZENKER–ZENKER 1960, idézi PABST–SUNDBERG 1992), és ezért igyekeznek lesüllyeszteni a gégét. A gége (nyugalmi helyzethez képest) alacsonyabb vertikális pozíciójának eléréséhez pedig SUNDBERG szerint hatékony indirekt eszköz lehet a hangzók kerekítéses ejtése (szíves szóbeli közlés). Valóban, az ajakkerekítéses hangzók (pl. az /y:/) esetében a gége jóval (a nyugalmi pozíciónál) lejjebb helyezkedik el, mint az azonos nyelvállásfokú ajakréses párjuk (pl. az /i:/) esetében, ahol a gége a nyugalmi pozíciónál magasabbra emelkedik (lásd pl. HOOLE–KROOS 1998).

5.2.2. Az ejtési szándéknak megfelelő magánhangzó-azonosítás az értelmes és az értelmetlen szavakban

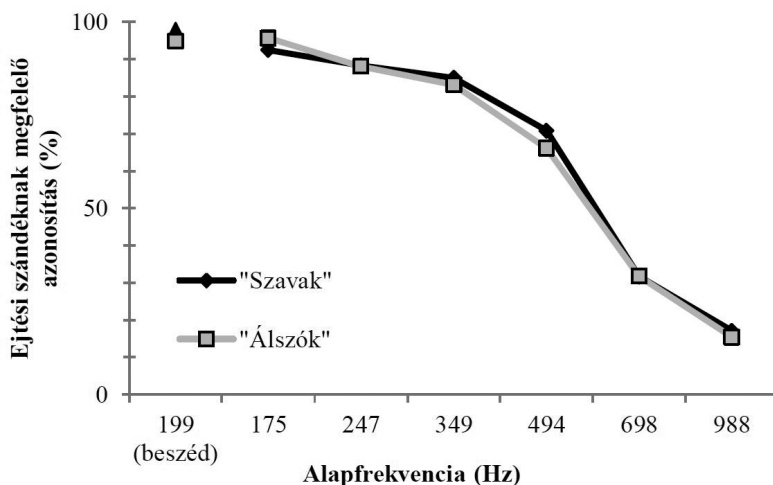
A 22. ábra az ejtési szándéknak megfelelő válaszok az egyes alapfrekvenciánként jellemző összesített arányát mutatja a „nem énekes” résztvevőknél. Az ábra a következőképpen olvassandó: míg a „nem énekes” adatközlőknél 494 Hz alapfrekvencián a „Szavak” kondícióban a (20 adatközlő \times 6 magánhangzó \times 2 ismétlés =) 240 magánhangzóválasz 70%-a egyezett meg az ejtési szándéknak megfelelő magánhangzóval, addig az „Álszók” kondícióban a 240 magánhangzó-válasznak csak 60%-a volt az ejtési szándéknak megfelelő. Az első, a beszédre kapott észlelési adatokat mutató adatpontot itt is elkülönítve jelenítettük meg, hiszen ez nem csak alapfrekvenciájában, de a produkció módjában is eltért a többi adatunktól.

Az ábrán jól látható, hogy a „nem énekesek” esetében a magánhangzó-azonosítás tendenciái gyakorlatilag megegyeztek a két kondíció esetében. Az apróbb eltérések ellenére a két adatsor közti különbség a páros t -próba tanúsága szerint statisztikailag sem szignifikáns. A legmagasabb (95% körüli) azonosítási arányokat mindkét kondícióban a beszéd módban látjuk. Ehhez képest az összes, még a beszéd alapfrekvenciájához viszonyítva alacsonyabb alapfrekvenciájú énekelt alapfrekvencia esetében is alacsonyabbak a százalékos arányok. Az általános tendenciák alapján azt mondhatjuk, hogy az ejtési szándék szerinti azonosítás aránya mindkét kondícióban csökkent az alapfrekvencia emelésével, mégpedig fokozatosan, ugyanakkor 500 Hz (azaz a h' zenei hang) környékén hirtelen esés volt tapasztalható. A két kondíció egyetlen ponton mutat szembevetőbb különbségeket: 494 Hz alapfrekvencián (a h' zenei alaphangon) az adatközlők teljesítménye 10%-kal alacsonyabb az álszókban, mint a szavakban.



22. ábra: Az egyes alapfrekvenciákon bemutatott összes stimulus ejtési szándék szerinti azonosítási aránya a két kondícióban a „nem énekes” adatközlőknél

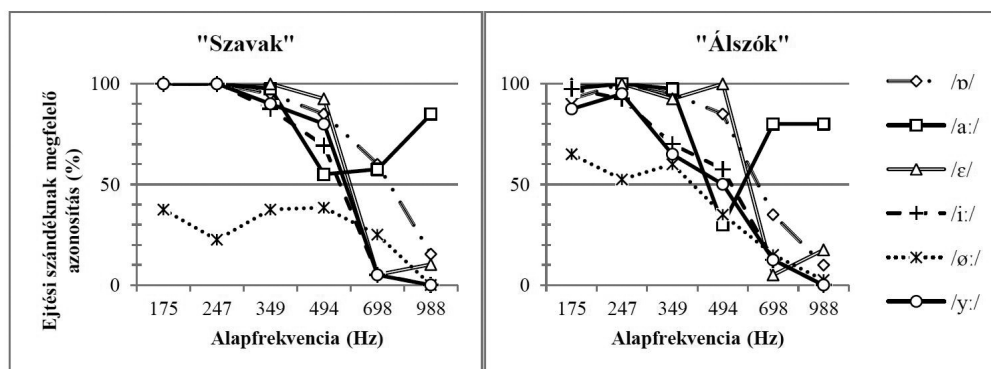
A 23. ábrán az „énekes” résztvevőknél tapasztalt azonosítási arányok láthatók. Az előzőekhez hasonlóan itt is azt találtuk, hogy a két kondícióban kapott válaszok aránya nem tér el egymástól statisztikailag igazolható módon (páros t -próba). Az azonosításban mutatott magasabb arányokat az adatközlők e csoportjában is a beszéd módban tapasztaltuk. Érdekes módon az „énekesek” esetében nem tapasztaltunk még olyan apróbb különbségeket sem a kondíciók között, mint a „nem énekesek”-nél.



23. ábra: Az egyes alapfrekvenciákon bemutatott összes stimulus ejtési szándék szerinti azonosítási aránya a két kondícióban az „énekes” adatközlőknél

5.2.3. A magánhangzók ejtési szándéknak megfelelő azonosítása a két kondícióban

A 24. ábrán az ejtési szándéknak megfelelő azonosítási arányokat látjuk magánhangzó-minőségenkénti bontásban, két kondícióban (balra: „Szavak”, jobbra: „Álszók”) a „nem énekes” adatközlőknél. Az eredmények összehasonlítására ismételt méréses vairanciaanalízist alkalmaztunk a „kondíció” faktorról („Szavak”, „Álszók”), valamint a „magánhangzó-minőség” faktorról. A vizsgálat szerint nem mutatható ki a „kondíció” faktor főhatása a magánhangzó-minőségek figyelembevételével sem.



24. ábra: Az ejtési szándék szerinti magánhangzó-azonosítás tendenciái a „Szavak” (balra) és az „Álszók” (jobbra) kondícióban a „nem énekes” csoportban

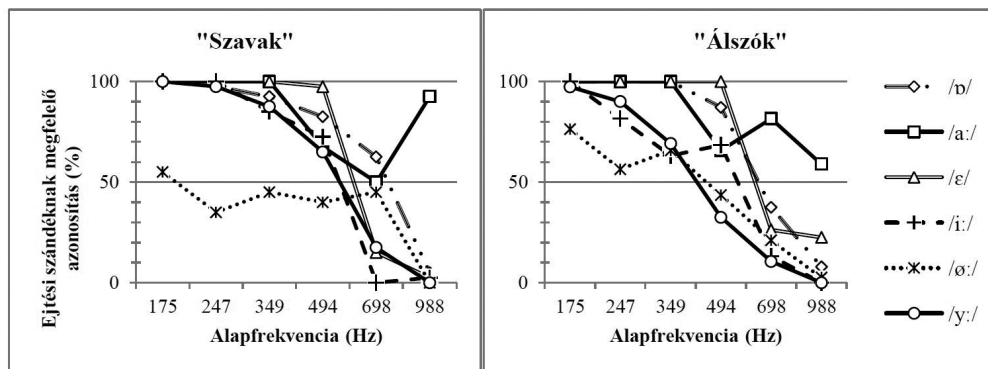
Annak ellenére, hogy a „Szavak” kondícióban az egyes ejteni szándékozott magánhangzó-minőségek eltalálására akár több esélye is lehetett az adatközlőknek, az /ø:/ esetében minden alapfrekvencián nagyon alacsony, illetve az „Álszók” kondícióban tapasztalhatóan általánosan is alacsonyabb azonosítási arányokat találtunk még az összesített eredmények alapján kritikusabbnak tűnő 698 Hz alapfrekvencia (f'') alatt is. Az /ø:/ azonosítása a beszédben is elmarad a többi magánhangzótól, ugyanis csak 95%-os, szemben a többi vokálisra jellemző 100%-kal (az ábrán ezek a beszéd módra kapott eredmények nem szerepelnek). A magánhangzónkénti összevetésből az is kiderült, hogy a 22. és 23. ábrán látott, a beszédre jellemző alacsonyabb (nem 100%-os) azonosítási arányok az /ø:/ bizonytalanak tűnő hangszínéből fakadnak.

A legnagyobb biztonsággal az ejtési szándék szerint azonosított magánhangzó mindkét kondícióban az /a:/ volt, mindazzal együtt, hogy egy ponton, 494 Hz alapfrekvencián hirtelen lecsökkent az azonosítási arány, és csak az f_0 további emelésével emelkedett meg ismét. Bár az ejtési szándék szerinti azonosítás arányaiban talált esés jelentősebbnek tűnik az „Álszók” kondícióban, a „Szavak” kondícióban a százalékok a sorrendben következő, azaz 698 Hz alapfrekvenciájú hangmagasságon is alacsonyabbak maradtak.

A „Szavak” és „Álszók” tendenciái között talált egy kisebb különbség az elől képzett első nyíltsági fokú /i:/ és /y:/ eltérő viselkedése a két kondícióban kritikusnak tűnő 698 Hz alatt (ahol az ejtési szándék szerinti azonosítás még általánosságban magasabb arányokat mutat). Az /i:/ és /y:/ ugyanis a „Szavak” kondícióban (tehát a *szír* és *szűr* szavakban) valamivel magasabb arányban értelmeződött az ejtési szándéknak megfelelően, mint az „Álszók” kondícióban (azaz a *dír* és *dűr* hangsorokban). A két kondíció között ebben a tekintetben a legnagyobb különbséget az /y:/-nél tapasztaltuk, melynek az ejtési szándék szerinti azonosítása 494 Hz alapfrekvencián 30%-kal kisebb volt az „Álszók” esetében.

A 25. ábra összegzi az énekes adatközlők eredményeit. Ezekre az adatokra (megint csak ismételt méréses varianciaelemzéssel végezve az összehasonlítást) szintén azt találtuk, hogy

nem mutatható ki a „kondíció” főhatás a magánhangzó-minőségek figyelembevételével sem. Mind az énekesekre, mind pedig a nem énekes résztvevőkre kapott eredmények azt jelzik, hogy bár az egyes magánhangzókra kapott adatok többé-kevésbé eltérhetnek az egyes alaphangfrekvenciákon, de a „kondíció” tekintetében nem különülnek el szisztematikusan, ha az összes alaphangfrekvenciát figyelembe vesszük az összehasonlításban.

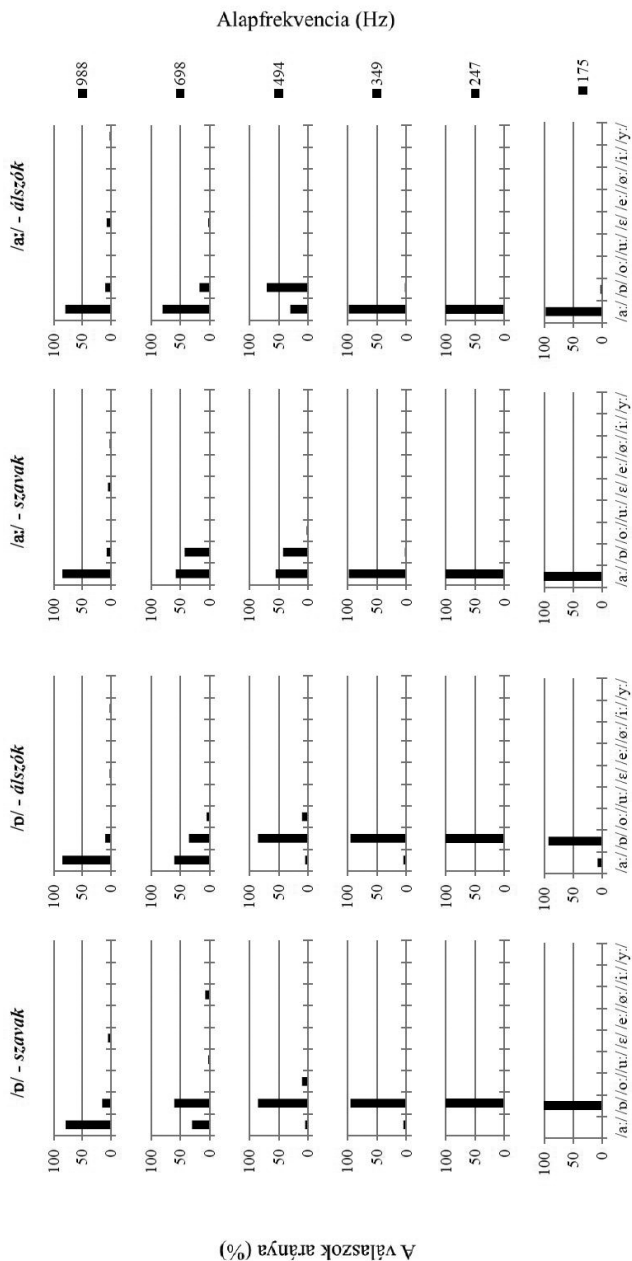


25. ábra: A magánhangzó-azonosítás tendenciái a „Szavak” (balra) és az „Álszók” (jobbra) kondícióban az „énekes” csoportban

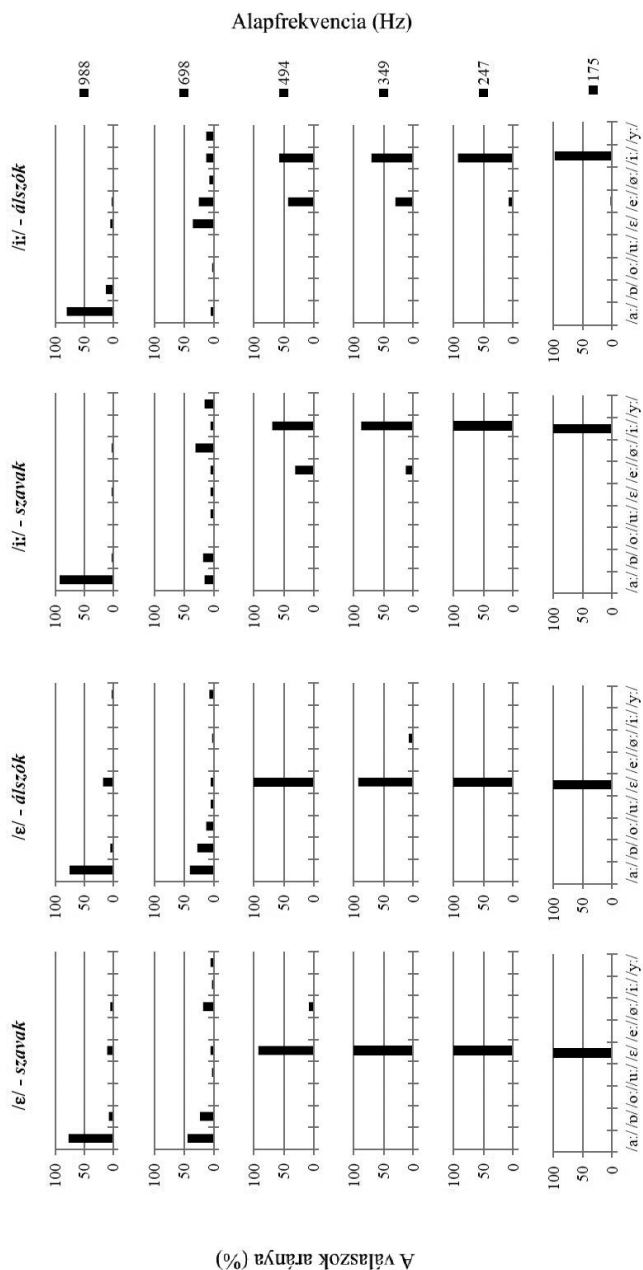
Az /ø:/ sajátosságos viselkedése, azaz a szavakban és álszókban tapasztalt azonosítási tendenciáinak eltérése az „énekes” csoportban is megfigyelhető volt, ugyanúgy, ahogy az /a:/ magas arányú ejtési szándék szerinti azonosítása. Eltérés azonban, hogy az „énekes” csoportban nem látunk olyan drámai esést az /a:/ esetében az „Álszók” kondícióban 494 és 698 Hz környékén.

5.2.4. A tévesztések tendenciái az egyes magánhangzók esetében az alaphangfrekvencia függvényében

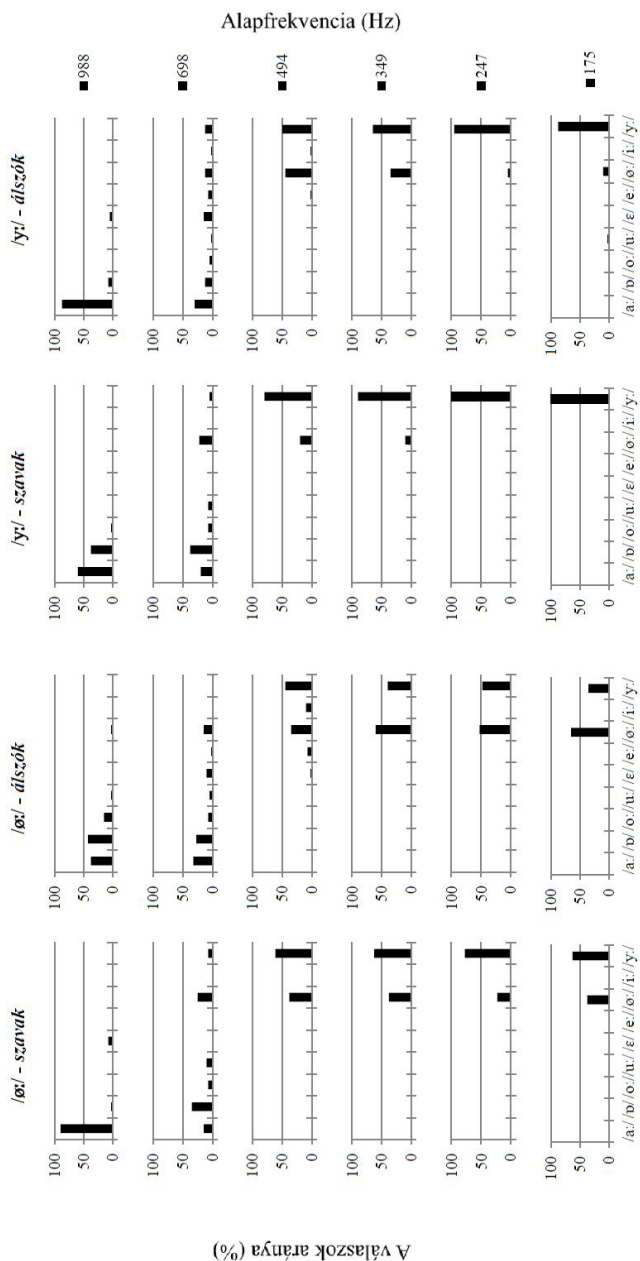
A 26. a), 26. b) és 26. c) ábrán a „nem énekes” csoportnak az egyes ejteni szándékozott magánhangzókra érkezett válaszait látjuk az alaphang függvényében a két kondícióban. Minden panel az egy-egy alaphangfrekvencián kapott összes választ (és a kilenc válaszlehetőséget) mutatja (beleértve az ejtési szándéknak megfelelő hangokra kapott válaszokat is) úgy, hogy (a hangmagasságérzetnek megfelelő sorrendben) a legalsó panel a legmélyebb, a legfelső pedig a legmagasabb alaphangfrekvenciára kapott válaszokat reprezentálja. (Ezek az ábrák a beszédmintákra kapott válaszokat nem tüntették fel.)



26. a) ábra: Az /b/ és /a:/ magánhangzókra adott válaszok (százalékos arányban kifejezett) alakulása a két kondícióban az f₀ függvényében a „nem énekes” adatközlőknél



26. b) ábra: Az /e/ és /i:/ magánhangzókra adott válaszok (százalékos arányban kifejezett) alakulása a két kondícióban az f_0 függvényében a „nem énekes” adatközlőknél



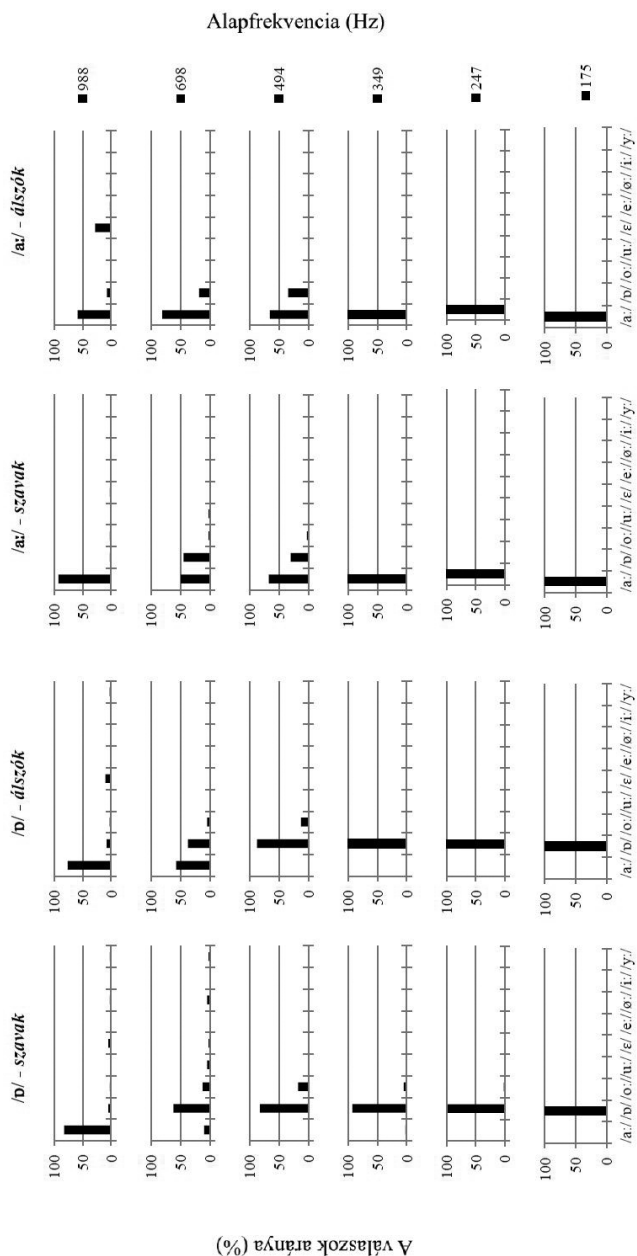
26. c) ábra: Az /ø:/ és /y:/ magánhangzókra adott válaszok (százalékos arányban kifejezett) alakulása a két kondícióban az f_0 függvényében a „nem énekes” adatközlőknél

Az előzőekben azt láttuk, hogy a „nem énekesek” mindkét kondícióban az /ø:/ hangot azonosították a legkisebb arányban az ejtési szándék szerint. Az ábra szerint ennek az oka az /ø:/-nek /y:/-ként történt nagyarányú azonosítása. Az /i:/ és /y:/ esetében magasabb ejtési szándék szerinti azonosítási arányokat találtunk a „Szavak” esetében. A 26., és különösen a 26. b) ábra tanúsága szerint ez az /i:/ esetében feltehetőleg arra vezethető vissza, hogy míg az „Álszók” esetében 247 és 494 Hz között az /i:/-re egyre nagyobb számban érkeztek /e:/ válaszok, a „Szavak” esetében az /e:/ válaszok megszorodása nem olyan nagy mértékű, hiszen az /e:/ válasz a „Szavak” esetében elvileg illegitim volt. Különösen érdekes megfigyelni, hogy a „tiltás” ellenére mégis milyen nagy arányban jelenik meg az /e:/ válasz a „Szavak” kondícióban is. Az /y:/ esetében a két kondíció közti különbséget 698 Hz alatt egyértelműen az /ø:/ válaszok arányának eltérése okozza (ez mindkét kondícióban legitim válasznak minősült). Az adatokból az is kiderül, hogy az egyébként általánosan nagy arányban az ejtési szándéknak megfelelően azonosított /a:/-ra az „Álszók” kondícióban (és főként csak 494 Hz-en) több /b/ válasz érkezett, mint a „Szavak” kondícióban.

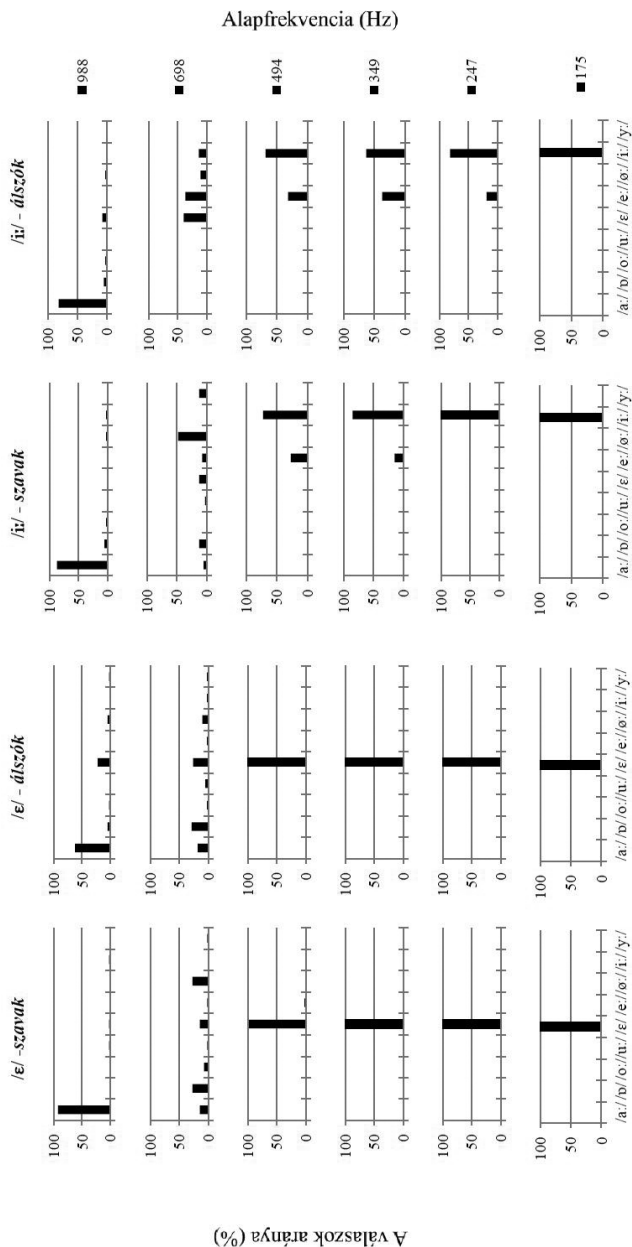
Az általános azonosítási tendenciákat összefoglaló 22. ábra alapján (mely 698 Hz alaphangfrecvenciától hirtelen esést mutatott) az összes választ bemutató 26. ábra figyelembevételével megállapítható, hogy a zenei *f* hangon, azaz 698 Hz alaphangfrecvencián a legnagyobb nyíltsági fokú /b/ és /a:/ magánhangzók kivételével minden vizsgált magánhangzóra közel a véletlennek megfelelő százalékban érkeztek válaszok kondíciótól függetlenül (bár sok esetben az /b/ és /a:/ válaszok 30% körüli dominanciájával). A zenei *h* hang, azaz 988 Hz alaphangfrecvencián azonban nagyon erős tendenciaként mutatkozott meg az /b/ és /a:/ válaszok megnövekedett aránya. A 26. ábrán látható adatok ezen túlmenően azt is megvilágítják, hogy a „Szavak” és „Álszók” kondíciók közti átlagosan legnagyobb, azaz 10%-nyi különbség (melyet 494 Hz alaphangfrecvencián tapasztaltunk) legnagyobb részt az /y:/-re kapott /ø:/ válaszoknak (rendre 11% és 19%), kisebb részben az /a:/-ra kapott /b/ válaszoknak (rendre 24% és 29%), legkisebb részben pedig az /i:/-re kapott /e:/ válaszoknak (rendre 17% és 18%) köszönhető. A két kondíció közti igen kis különbséget tehát összesen három magánhangzó észlelési tendenciáinak még apróbb eltérései okozzák.

A 27. ábra az „énekes” csoportban kapott válaszokat összesíti alaphangfrecvenciák, magánhangzók és kondíciók szerinti bontásban. Ebben a csoportban nagy vonalakban az előzőekben bemutatottakhoz hasonló tendenciákat láttunk. Az „énekesek” egy, a kísérletet követő jellemző megjegyzése szerint az /i:/ és /e:/ ejtése magas alaphangfrecvencián egyes énekeseknél gyakran inkább /y:/- és /ø:/-szerű realizációkat jelent (mely „csalásra” explicit énektanári instrukciók vezetnek az énekeseket), ezért az elemzés legelején felmerült, hogy az „énekesek” esetében ezeknél a hangzóknál a „nem énekesek”-nél tapasztalhatóktól eltérő tendenciákat várhatunk. A megjegyzésből egyrészt következtethetnénk az /y:/ és /ø:/ válaszok preferenciájára, másrészt pedig akár éppen az /i:/ és /e:/ válaszok preferenciájára is, ha azt feltételezzük, hogy az „énekesek” igyekeztek „visszakövetkeztetni” a magánhangzók eredetinek vélt minőségére, tehát kompenzálták a vélt hanghelyettesítésen alapuló produkciós stratégiát. A 27. ábrán bemutatott adatok tanúsága szerint azonban egyik tendencia sem volt jellemzőbb

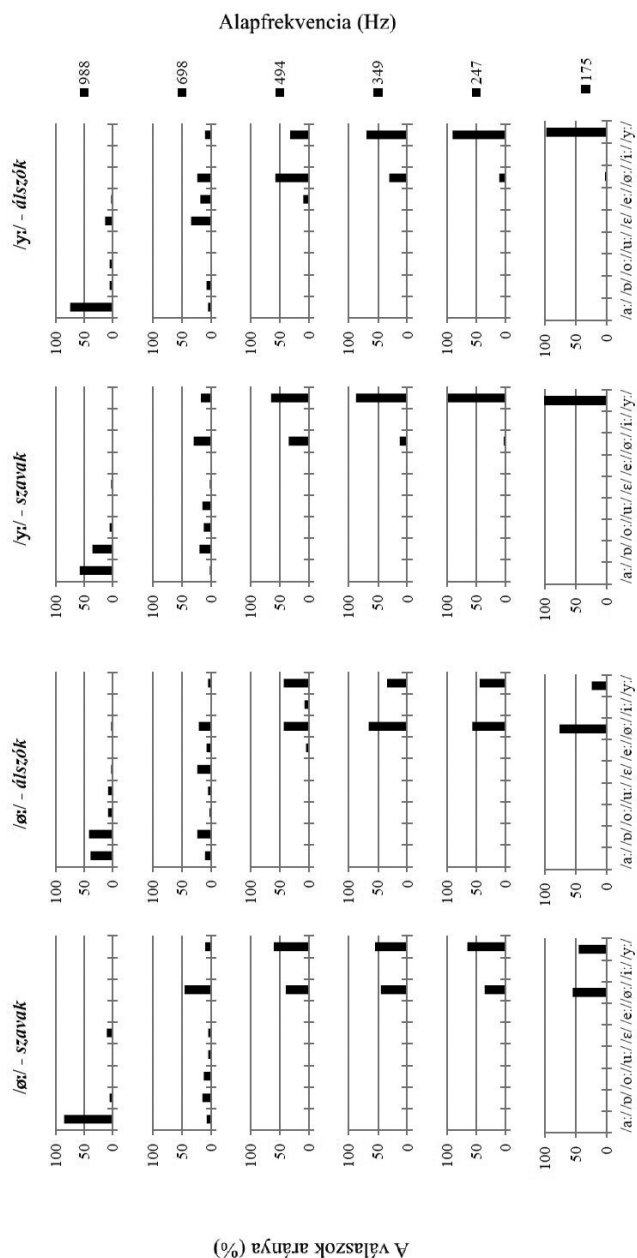
az „énekesek”-re, mint a „nem énekesek”-re. Ugyanúgy, ahogy a „nem énekesek”-nél láttuk, az „énekesek”-nél is jellemzően az /ø:/-re adott válaszként láttunk nagyobb számban /y:/-t (a „Szavak”-ban és az „Álszók”-ban rendre 62 és 45%), míg az /ø:/ is elsősorban az /y:/-re adott válaszként jelent meg leginkább 494 Hz alapprofrekvencián. A „nem énekes” csoporttól eltérően az „énekesek”-nél az /i:/-re adott /e:/ válaszok aránya a „Szavak” kondícióban is magas (mindkét kondícióban 25% körüli) volt, azaz ők nagyobb számban választották az illegitim alakot, mint a „nem énekesek”. Ugyanakkor az is igaz, hogy az „énekesek”-nél 698 Hz alapprofrekvencián a „Szavak” kondícióban lecsökkent az /e:/ válaszok aránya, és megnövekedett az /ø:/ válaszok aránya (48%), míg az „Álszók” kondícióban továbbra is az /ε/ és /e:/ dominált (rendre 39% és 37%). A „nem énekesek”-nél emellett megfigyelhettük azt is, hogy a zenei *f* hangon, azaz 698 Hz alapprofrekvencián az /v/ és /a:/ hangzók kivételével mind a hat vizsgált hangzóra nagyjából a véletlenhez közelítő arányban (de tendenciózusan az /v/ és /a:/ nagyobb arányával) jelentek meg válaszok. Ehhez képest eltérést az „énekesek”-nél csak az imént említett, /i:/-re kapott nagyarányú /ø:/, /ε/ és /e:/ válaszok jelentettek. A „nem énekesek” válaszaihoz hasonlóan az „énekesek” válaszairól is elmondható, hogy azok között a zenei *h* hangon, azaz 988 Hz alapprofrekvencián egyértelműen az /v/ és /a:/ hangokra „tévesztések” domináltak.



27. a) ábra: Az /b/ és /a:/ magánhangzóra adott válaszok (százalékos arányban kifejezett) alakulása a két kondícióban az f_0 függvényében az „énekes” adatközlőknél



27. b) ábra: Az /e/ és /i:/ magánhangzóra adott válaszok (százalékos arányban kifejezett) alakulása a két kondícióban az f_0 függvényében az „énekes” adatközlőknél



27. c) ábra: Az /ø/ és /y/ magánhangzóra adott válaszok (százalékos arányban kifejezett) alakulása a két kondícióban az f_0 függvényében az „énekes” adatközlőknél

5.3. Következtetések

A kísérletben az /ɒ a: e: i: ø: y:/ magyar magánhangzók azonosítását vizsgáltuk az éneklésben értelmes magyar szavakban és álszókbán, különös tekintettel a magas alapfrekvenciákra, laikus beszélők és énekesek csoportjában. SMITH és SCOTT (1980) értelmes szavakban és izoláltan ejtett magánhangzók azonosítását vizsgáló kutatásából, valamint saját (a 4. fejezetben bemutatott), értelmetlen szavakban és izoláltan ejtett magánhangzók azonosítását elemző vizsgálatunk eredményeiből a vizsgálat előtt azt a feltételezést fogalmaztuk meg, hogy nem hatékonyabb az ejtési szándéknak megfelelő magánhangzó-azonosítás az értelmes szavakban akkor, ha az értelmes és értelmetlen szavakban (azaz álszókbán) ejtett magánhangzók azonosítását nem jelentősen eltérő (viszont akár nem is szükségszerűen ugyanannyi), nagyszámú válaszlehetőséget megengedő feladatokban vetjük össze. A kísérlet egyik célja ennek a feltételezésnek a vizsgálata volt. A kísérlet másik célja énekes és nem énekes hallgatók teljesítményének összehasonlítása volt a magánhangzó-észlelés tekintetében, abból az intuitív és részben a beszédészlelés motoros elméletére (LIBERMAN–MATTINGLY 1985) is támaszkodó kérdésfelvetésből kiindulva, mely azt firtatja, hogy vajon az énekesek produkciós gyakorlottsága az énekelt magánhangzók azonosítására is hatással van-e. Előzetes feltevésünk szerint az énekesek nagyobb arányban képesek az ejtési szándéknak megfelelő hangzóazonosításra, mint a nem énekes hallgatók.

A kísérlet eredményei az első feltételezést alátámasztották, míg a másodikat cáfolták. Azt találtuk ugyanis, hogy sem a két vizsgált kondíció (értelmes szavak és álszók), sem pedig a két adatközlőcsoport (a nem énekesek és az énekesek) között nem mutathatók ki jelentős eltérések. Más szóval a bottom-up percepciós folyamatok aktivációja (azaz az értelmes, létező szó mint kontextus) kísérleti eredményeink szerint nem segítette jobban az észlelőket a magánhangzó ejtési szándék szerinti azonosításában akkor, ha az értelmes szavak az értelmetlen szavakhoz közel azonos számú (tehát még csak nem is teljesen egyező valószínűségű) választ engedtek meg. Megállapítottuk, hogy a legnagyobb különbség (ez mindösszesen 10% volt) a két kondíció között az /y/-re kapott /ø:/ válaszok, valamint kisebb részben az /a/-ra kapott /ɒ/ válaszok és az /i/-re kapott /e:/ válaszok kondícióként eltérő számának volt köszönhető (és ezt kizárólag a nem énekes adatközlőknél tapasztaltuk). Ezek az eltérések az egyes magánhangzókra vetítve is igen kicsinek bizonyultak. Az első hibatípus feltehetően (legalábbis részben) az /ø:/ egyébként is megmutatkozott bizonytalan percepciós minőségével hozható összefüggésbe (amelyet több ponton, így a beszéd módra kapott eredményekben is tapasztaltunk a teszt során). Többek közt a beszéd módban talált alacsonyabb azonosítási arányokból is kiindulva feltételezzük, hogy a kísérletben részt vevő szoprán énekes adatközlő /ø:/ beszédhangjának realizációi az adatközlők perceptuális /ø:/ kategóriájának kevésbé prototipikus tagjai voltak. Ezt a feltételezést az adatközlőknek az értelmes szavas, tehát az elsőként elvégzett tesztben mutatott teljesítménye is alátámasztja, melyben az /ø:/ ejtési szándéknak megfelelő azonosítása általánosan nagyon alacsony volt.

Értelmezésünk szerint a két kondíció közti, az /ø:/ tekintetében látható különbség, azaz az /ø:/ valamivel magasabb arányú, az ejtési szándéknak megfelelő azonosítása az értelmetlen szavakat tartalmazó, másodikként elvégzett tesztben hiperkorrekció eredménye lehetett. Feltételezhetjük ugyanis, hogy az első részben az adatközlők érzékelték, hogy meglehetősen kevés egyértelműen /ø:/-szerű hangot, ellenben sok, az /ø:/ és /y:/ kategóriák határán álló hangot hallottak, és hogy ezzel együtt felismerték azt is, hogy a teszt kitöltésekor ezekre a bizonytalanabb besorolású hangokra összességében viszonylag kevés /ø:/ választ adtak. Emiatt pedig azt is valószínűsíthetjük, hogy az adatközlők a második feladatban ezt a válaszadási tendenciát kompenzálni akarták, tehát ebben a feladatban már szándékosan inkább /ø:/-t jelöltek nagyobb arányban válaszként a kérdéses esetekben (tehát /y:/ vagy /ø:/ elhangzásakor).

Az /ɒ/ válaszok alacsonyabb aránya a „Szavak” kondícióban feltehetően (legalábbis részben) magyarázható azzal, hogy az adatközlők némileg feszélyezve érezhették magukat az /ɒ/ hangzó megjelölésével kapcsolatban a *sVr* (értelmes szavas) környezetben. Emellett az sem kizárt, hogy azt feltételezték, hogy a feladat becsapós, és ezt próbálták – ismét csak – túlkorrigálni.

Végezetül az /i:/ megvalósulásaira kapott /e:/ válaszok, illetve azok arányának eltérése a két kondícióban – mint fentebb említettük – azért is lehet valószínű, mert az /e:/ elméletileg (az adatközlőknek adott instrukciók szerint) illegitim válasz volt az értelmes szavas feladatban. Újra kiemelendő itt az az érdekes megfigyelés is, mely szerint az /e:/ ilyen tekintetben illegitim státusza ellenére is gyakori válasz volt az első (értelmes szavakat tartalmazó) feladatban, ami feltehetőleg az /i:/-t érintő erős hangszínváltozásnak volt köszönhető – az /i:/ minősége ugyanis korábbi vizsgálataink (lásd 4. fejezet) szerint már 349 Hz alaphangfrekvencia környékén is megváltozik, és a megváltozott minőségű hangzót ezen az alaphangfrekvencia-tartományon még leginkább /e/-ként ismerik fel a hallgatók.

Mind az /y:/-re kapott /ø:/ válaszok, mind pedig az /i:/-re kapott /e:/ válaszok azokkal a korábbi vizsgálati eredményekkel csengnek egybe, melyek a produkciós és percepciós működések közti egyszerűbb, lineáris összefüggéseket feltételeztek, és emiatt úgy vélték, az alaphangfrekvencia emelésével tapasztalható produkciós nyitódást percepciós nyitódásnak kell kísérnie (lásd SCOTTO DI CARLO–GERMAIN 1985, HOLLIEN et al. 2000). Ugyanezen eredményeknek és feltételezéseknek megfelelő tendencia az /a:/ nagyarányú azonosítása is a teljes alaphangfrekvencia-tartományon – hiszen az /a:/ magas első formánása miatt az alaphangfrekvencia emelésével akár magasabb f_0 -értékekig is „intakt” maradhat az $F_1 : f_0$ formánshangolástól –; valamint az /ɒ/ és /a:/ válaszok megnövekedett száma 600 Hz alaphangfrekvencia fölött (de különösen 988 Hz alaphangfrekvencián).

Az alaphangfrekvencia emelésével általános tendenciaként megjelent a válaszok egyre nagyobb szóródása, mely a zenei f'' hangon, azaz 698 Hz alaphangfrekvencián – a legnyíltabb /ɒ/ és /a:/ magánhangzók kivételével – minden vizsgált magánhangzó esetében a véletlenszerűnek megfelelő válaszadásig fokozódott (kondíciótól és résztvevőcsoporttól függetlenül). Ebből – a korábbi, 2., 3. és 4. fejezetekben is ismertetett vizsgálatainkkal egybehangzó módon – arra

következtethetünk, hogy a magánhangzókat az alapfrekvencia emelésével egyre inkább jellemző fonetikai alulspecifikáltság (azaz az artikulációs változtatások és a spektrális alul-min-tavételezettség) következtében a zenei f'' hang magasságának környékén az egyes magánhangzók minősége gyakorlatilag teljesen elbizonytalanodik. A jelen kötetben bemutatott vizsgálatok kísérleti anyagának korlátait is tekintetbe véve tehát levonhatjuk azt a következtetést, hogy az egyes magánhangzók a zenei f'' hang magasságának környékén kerülnek a legtávolabb az egyes prototipikus (a magyar magánhangzókra jellemző) percepciók kategóriáktól négy olyan énekes vizsgálati személy produkciójában, akik bizonyítottan tekinthető módon alkalmazzák az $F_1 : f_0$ hangolásnak a professzionális szoprán énekeseknél feltételezett rezonanciastratégiáját (lásd 2., 6. és 7. fejezet).

Eredményeink szerint az imént említett zenei hangmagasság fölött, mégpedig a h'' zenei hang, azaz 988 Hz alapfrekvencia magasságában, bizonyos értelemben, ismét megerősödik a prototípus hatás. Itt ugyanis a kísérletben (gyakorlatilag az ejtési szándék szerinti magánhangzó-minőségtől függetlenül) kimagaslóan megnövekedett az / υ / és / a / válaszok aránya a válaszok között. Ezt a következőképpen magyarázhatjuk. 988 Hz alapfrekvencián a magánhangzó-észlelés szempontjából legfontosabbnak vélt 5000 Hz alatt bármely magánhangzó csak a 988 Hz (f_0), 1976 Hz ($2f_0$), 2964 Hz ($3f_0$), 3952 Hz ($4f_0$) és a 4940 Hz ($5f_0$) frekvenciákon tartalmazhat intenzív összetevőket, hiszen csak ezek a frekvenciakomponensek jelennek meg egy 988 Hz alapfrekvenciájú (kvázi)periodikus komplex hang spektrumában, és ezeket erősítheti (vagy sem) egy-egy üregi rezonancia. Ezek közül az első két harmonikus, azaz az f_0 és a $2f_0$ éppen a női ejtésben az / a / F_1 és F_2 értékéhez közeli frekvenciaértéken helyezkedik el. Ha tehát ezek a harmonikusok erősítettek a hangzó spektrumában (az f_0 erősítettségét az $F_1 : f_0$ hangolás miatt pedig feltételezhetjük), akkor az adott hangzó spektrális szerkezetét tekintve a női ejtésben megvalósuló / a / magánhangzóhoz hasonlít a legjobban.

A zenei f'' hangon (698 Hz alapfrekvencián) erősen elbizonytalanodó, majd a zenei h'' hangon (988 Hz alapfrekvencián) ismét valamivel egyértelműbbnek látszó hangzóminőségek azt jelzik, hogy bár az éneklésben az alapfrekvencia emelésével és a fokozatos produkciós nyílással párhuzamosan nem kategorikusan változnak az egyes magánhangzó-minőségek, de kellően magas (1000 Hz körüli) alapfrekvencián mégis tapasztalható kategoriális váltás. Itt ugyanis – feltehetően a felharmonikusok frekvenciaértéke miatt – a megvalósuló magánhangzók zöme a legnyíltabb / a /-ként (esetleg / υ /-ként) azonosítható.

A nem énekes és énekes adatközlők percepcióját érintően azt láttuk, hogy az ejtési szándéknak megfelelő magánhangzó-azonosítás akkor sem volt magasabb, ha az észlelő maga is énekes, tehát ha a résztvevő artikulációs gyakorlattal, illetve explicit vagy implicit tudással rendelkezett az éneklés artikulációs sajátosságairól. Az énekesek és nem énekesek azonosítási tendenciáiról ezen felül azt is megállapíthattuk, hogy egyéb módon (például a jellemző hibák tekintetében) sem tértek el szisztematikusan.

A kísérlet második feltételezésének cáfolatát a motoros elmélettel kapcsolatos ellenbizonyítékként értelmezhetjük akkor, ha a motoros elméletből következőnek tekintjük azt az állítást, mely szerint a szenzomotoros kapcsolatok révén a produkciós gyakorlottság vagy tapasztal-

talat percepcióis előnyt is jelent(het). Eredményeink ugyanis azt jelzik, hogy a rájuk jellemző artikulációs gyakorlottság ellenére az énekesek nem látszanak percepcióis előny birtokában lenni a nem énekesekhez képest. Ugyanakkor nem zárjuk ki annak a lehetőségét sem, hogy a motoros elméletből a jelen vizsgálati kérdésre más feltételezések is megfogalmazhatók volnának, melyekkel kapcsolatban eredményeink státusza nem feltétlenül egyértelmű. Ilyen feltételezés lehet például az, hogy a szenzomotoros kapcsolatok megléte és a motoros területek aktivitása a percepcióban valójában nem előnyt, hanem „csak” bizonyos tendenciák megjelenését jelenti. Ezzel kapcsolatos konklúziókra a jelen vizsgálat eredményeiből már csak sokkal óvatosabban vállalkozhatnánk, hiszen bár a két vizsgált csoport észlelési tendenciái között itt nem találtunk szisztematikus eltéréseket, ezek létét mégsem zárhatjuk ki teljesen. Ezt a kérdést egy alapvetően más jellegű (elsősorban csak az akusztikai elemzésen alapuló, értelmetlen szavakat vagy izolált ejtésű hangzók felismerését elemző) kísérletben gondoljuk vizsgálhatónak.

Első feltételezésünk teljesülését annak a felvetésnek a (bizonyos tekintetben közvetettnek tekinthető) bizonyításaként értelmezzük, hogy a SMITH és SCOTT (1980) vizsgálatában kapott eredmények, azaz a magas alaphfrekvencián énekelt magánhangzók ejtési szándéknak megfelelő azonosításának magas aránya mássalhangzó-környezetben valójában nem a fonetikai kontextus hagyományos (a koartikulációs formánsátmenetek mint másodlagos akusztikai kulcsok szerepét hangsúlyozó) értelmezésében jelent pozitív hatást – ahogyan azt a szerzők állítják –, hanem a kontextus fogalmának egy másik lehetséges interpretációjában. Ennek nyomán úgy látjuk, hogy a jelen vizsgálat eredményei egy, a kísérlet kérdésfelvetésénél meszszebbre vezető konklúziót is maguk után vonnak.

Mivel eddig már számos artikulációs és akusztikai vizsgálat bizonyította azt, hogy az énekelt magánhangzók artikulációs és akusztikai célja az alaphfrekvencia emelésével fokozatosan változik, a mássalhangzó-környezetben mint formánsátmenetekben kulcsolt magánhangzó-információról is csak azt feltételezhetjük, hogy erről a megváltozott artikulációval és akusztikai szerkezettel megvalósult magánhangzóról, nem pedig az ejteni szándékozott magánhangzóról szolgáltatathat információkat. Mivel azonban általában nem (és a jelen vizsgálatban sem) tapasztalunk egyértelmű kategóriaugrásokat a percepcióis tendenciákban, továbbá a szonoráns formánsátmenetekről (a magánhangzókhoz hasonlóan) is feltételezhetjük a spektrális alul-mintavételezettséget, a kontextus illetően hatásával kapcsolatban nem szűrhetünk le egyértelmű következtetéseket. Úgy véljük azonban, hogy egyértelműen elvethető az a feltételezés, melyet az éneklés szakirodalma nem ritkán sugall, és amely szerint a mássalhangzó-környezet a dinamikus akusztikai kulcsok, azaz a formánsátmenetek révén segítheti az ejteni szándékozott magánhangzók azonosítását. Mint említettük ugyanis, ezek a kulcsok inkább csak a megváltozott minőségű hangzók akusztikai kulcsolását biztosíthatnák.

A mássalhangzó-környezet mint értelmes szavakba ágyazottság ezzel szemben úgy tűnik, hogy elősegítheti a magas alaphfrekvencián énekelt magánhangzók az ejtési szándék szerinti azonosítását, ám pusztán azért, mert az értelmes szavak száma és a bennük előforduló magánhangzók minőségi különbségei növelik a „találat” valószínűségét ahhoz a felismerési

feladathoz képest, ahol az összes, az adott nyelvben létező magánhangzó lehetséges válasz lenne. Erre a következtetésre vezetnek SMITH és SCOTT (1980) eredményei is, legalábbis a jelen vizsgálat eredményeivel összevetve.

Megerősítettnek látjuk tehát azt a már a 4. és 5. fejezetekben is megfogalmazott egyik javaslatunkat, mely szerint a mássalhangzó-környezetnek tulajdonítható (a hangzóészlelést érintő) hatások vizsgálatában szükséges a „mássalhangzó-környezet” terminus differenciálása (az éneklés és a beszéd viszonylatában egyaránt). Ha ugyanis egyszerűen csak a „mássalhangzó-környezet” hatásairól teszünk említést, akkor félreérthetően utalhatunk mind a **koartikulációs formánsátmenetekre**, tehát azokra a beszédhangok környezetbe ágyazottságából eredő akusztikus kulcsokra, melyeket a beszédészlelés a beszédhangok azonosításában felhasználhat, mind pedig a beszédhangoknak az **értelmes szavakba ágyazottságára**, ami pedig a felismerendő beszédhangot tartalmazó hangsor lehetséges kiegészítéseinek száma révén gyakorol hatást a beszédpercepcióra (hiszen behatárolja, korlátozza a válaszlehetőségeket).

A két fent említett jelenség elválasztása többek között azért is gyümölcsöző lehet, mert segíthet feloldani az olyan jellegű ellentmondásokat (vagy inkább ellentmondásosnak látszó eredményeket), mint amilyeneket az énekelt magánhangzók azonosításával kapcsolatban is tapasztaltunk. A mássalhangzó-környezet itt ismertetett interpretációjának segítségével eredményeinket a következőképpen fogalmazhatjuk meg.

Eddigi eredményeink fényében azt állíthatjuk, hogy a magas alapfrekvencián énekelt magánhangzók esetében mind a magánhangzókat, mind pedig a szomszédos mássalhangzók és a magánhangzó közti koartikulációs formánsátmeneteket fonetikai alulspecifikáltság (azaz spektrális alul-mintavételezettség, illetve megváltozott artikuláció és akusztikum) jellemzi. Ebből kifolyólag jól magyarázható az, hogy az itt bemutatott vizsgálat tanúsága szerint a kontextus, még pontosabban a **koartikulációs formánsátmenetek** nem támogatják (hiszen nem is támogathatják) az ejtési szándéknak megfelelő azonosítást az éneklésben. Emellett azonban SMITH és SCOTT (1980) eredményeire támaszkodva azt is állíthatjuk, hogy a kontextus mint **értelmes szavakba ágyazottság** jelentős segítséget jelenthet a beszédészlelési rendszer számára az ejtési szándék szerinti hangzóazonosításban, hiszen ha a mássalhangzóváz kellően kisszámú kiegészítési lehetőséget enged meg, az megnöveli az egyes „megoldások”, így az ejtési szándék szerinti azonosítás valószínűségét is – az akusztikai információktól majdhogynem függetlenül.

Következésképpen úgy látjuk, hogy a magas alapfrekvencián énekelt magánhangzók azonosítása nem az elsődleges vagy másodlagos akusztikai kulcsok, tehát nem az akusztikai feldolgozás, hanem minden bizonnyal szinte kizárólag a **fonológia-fonotaktika**, valamint a **nyelvi kontextus**, azaz a szintaktikai, szemantikai és pragmatikai beágyazottság függvénye. A magas alapfrekvencián énekelt magánhangzók az ejtési szándéknak megfelelő azonosítása vélekedésünk szerint ugyanis elsősorban attól függ, hogy az azonosítandó beszédhango(ka)t tartalmazó szónak mekkora a szomszédsági sűrűsége (idegen szóval mekkora a szóhoz tartozó *phonological neighborhood density*), tehát mekkora azon fonológiai szom-

szédainak számossága, melyek csak egy (ráadásul a kérdéses) fonémában térnek el tőle. Másodsorban pedig nem szabad megfeledkeznünk arról sem, hogy egy-egy librettóban a magánhangzókat tartalmazó szavak általában tágabb szövegekörnyezetbe, azaz mondatba és hosszabb szövegbe is ágyazottak, ami a kérdéses szó (és magánhangzó) beazonosítását nyilvánvalóan szintén nagyban segíti.

6. Az énekelt magánhangzók akusztikai szerkezete néhány módszertani újítás tükrében

A 2. fejezetben bemutatott vizsgálatunkban a magas alapfrekvenciájú éneklésben megvalósuló magyar magánhangzókat vizsgáltuk akusztikai és percepciós szempontok szerint. Az akusztikai elemzés során kiemeltük, hogy mivel a magas alapfrekvencián a magánhangzó-realizációk formánsai, azaz a toldalékcso-re mint üregrendszerre az ejtés közben jellemző sajátrezonanciák nem állapíthatók meg egyértelműen az akusztikai jelből, az idézett fejezetben közölt méréseink eredményei a magas alapfrekvenciák esetében nem vonatkoztathatók közvetlenül az artikulációs beállításokra. Emlékeztetül, amint azt az 1.1.1.2. alfejezetben részletesen tárgyaltuk, a magas alapfrekvencián ejtett hangok esetében az alacsony alapfrekvencián megvalósított hangzókhoz képest kisebb sűrűségű a felharmonikus-szerkezet, tehát a hangzók spektrálisan alul-mintavételezetté válnak, ami az f_0 emelésével egyre fokozódik. Ennek következtében a magas alapfrekvenciák esetén az akusztikai jel Fourier-transzformáción alapuló elemzése kizárólag a spektrális energiamaximumokat teszi detektálhatóvá, nem pedig az azokat előidéző rezonanciákat. A rezonanciák frekvenciahelyeit ugyanis az egyes felharmonikusok erősítettsége mindössze csak sejteni engedi, azon keresztül, hogy az erősített felharmonikusokról valószínűsíthető, hogy azok beleesnek valamely rezonancia sávzélességébe. Mindennek következtében az akusztikai jel elemzése magas alapfrekvenciákon már nem enged megbízhatóan következtetni az ejtést jellemző artikulációs működésekre és akusztikai stratégiákra. Éppen ezért a szoprán éneklésben a magas alapfrekvencián feltételezett $F_1 : f_0$ hangolás pusztán az akusztikai kimeneti jel (azaz az énekhang) vizsgálatán keresztül nem bizonyítható maradéktalanul, a feltételezés ellenőrzéséhez olyan alternatív mérési eljárás szükséges, mely lehetővé teszi a toldalékcso sajátrezonanciáinak az alapfrekvenciától független mérését.

A jelen fejezetben közölt vizsgálat célja, hogy megállapítsa azokat az akusztikai változásokat, amelyek a magyar köznyelvi magánhangzók formánsszerkezetét jellemzik az alapfrekvencia emelésével a szopránéneklésben. Feltételezésünk szerint a magyar magánhangzóban is megjelenik a korábban, más nyelvek magánhangzóira leírt formánshangolási tendencia, azaz az első formánsnak az alapfrekvenciára hangolása ($F_1 < f_0$ helyzetben), de a hangolási tendencia megjelenésének „helye” (az alapfrekvencia emelésének folyamatára vetítve) a magánhangzó nyíltságától, illetve a magánhangzó beszédben megjelenő F_1 -értékétől függ.

Az akusztikai vizsgálatához két olyan alternatív mérési módszert alkalmazunk, amelyek korábbi vizsgálatok tanúsága szerint alkalmasak lehetnek arra, hogy segítségükkel még magas alapfrekvencián is megbízhatóan detektáljuk a toldalékcso-re jellemző sajátfrekvenciákat. Az egyik módszer a manuális inverz szűrés kiegészítve a hangszalagregzésre vonatkozó

információkkal: a szájüregből kisugárzott akusztikus jel mellett a hangszalagok működését is regisztráljuk elektrolottográf segítségével, és a formánsfrekvenciákat az akusztikai jel inverz szűrésével állapítjuk meg olyan módon, hogy a szűrésben az elektrolottogramban rejlő információkat is hasznosítjuk (vö. HERTEGÁRD–GAUFFIN 1993). A másik módszer pedig egy, a fonációtól független, külső gerjesztés használata: rögzítünk és elemzünk egy, a toldalékcso külső gerjesztése mellett felvett audiojelet is (vö. SUNDBERG 1975), így erősítve meg (és egészítve ki) az inverz szűrés során nyert akusztikai adatokat. (Az inverz szűrés és az elektrolottográfiát, valamint ezek felhasználását a formánsmérésekben a következő két, azaz a 6.1. és 6.2. *alfejezet*ben mutatjuk be; a toldalékcso „újra-mintavételezésének” metódusát, azaz a külső gerjesztés használtát pedig a 6.4. *alfejezet*ben ismertetjük.) A jelen vizsgálat újdonsága, hogy a két, korábban más vizsgálatokban már sikerrel alkalmazott módszert egymással elsőként ötvözve kíván a magas alapfrekvenciájú énekprodukción során jellemző toldalékcsovi rezonanciákról képet adni, miközben minden eddigi kísérletnél megbízhatóbb formánsadatokat szolgáltat a magyar köznyelvi magánhangzók megvalósulásáról a szopránéneklésben.

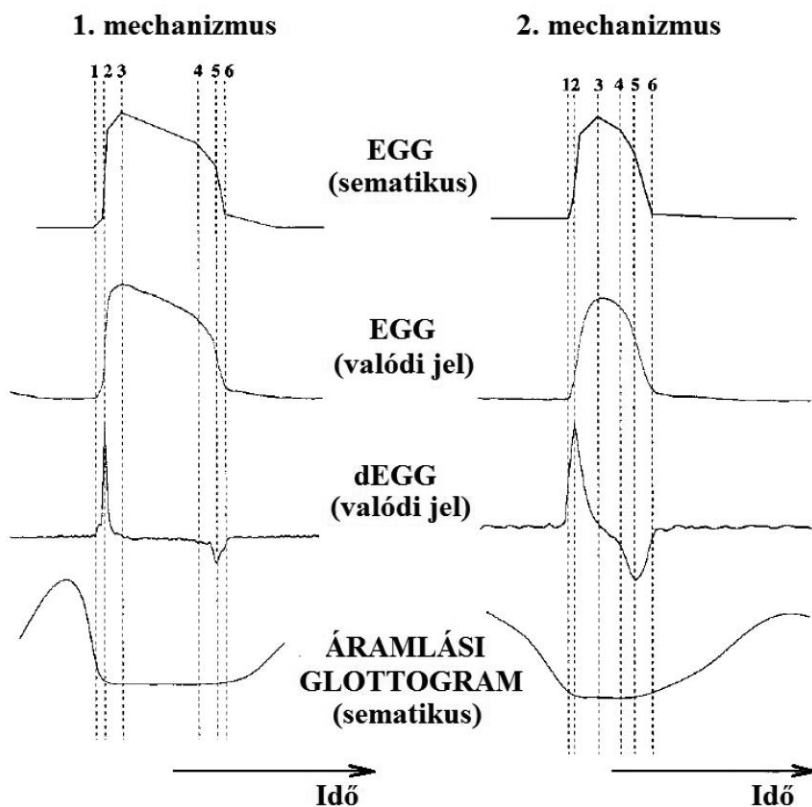
6.1. Az elektrolottográfia és az elektrolottogram

Az itt következőkben a vizsgálatban használt egyik technológiát, az elektrolottográfiát mutatjuk be olyan szempontok alapján, melyek megvilágítják, miért és hogyan használható ez az eljárás a formánsmeghatározásban.

Az elektrolottográf olyan (az egészségre teljesen ártalmatlan) eszköz, mely segítségével detektálható a hangszalagok zárt állása a zöngképzés során. Az elektrolottográf működési elve, hogy két elektróda segítségével kisfeszültségű, magas frekvenciájú elektromos áramot vezet át a pajzsporc egyik oldaláról a másikra a hangszalagok magasságában, és az átjutó jel feszültségváltozásait rögzíti. Ezt a jelet hívják elektrolottogramnak (EGG).²⁷ Mivel a hangszalagok közelítése és távolítása miatt az átjutó áram mennyisége a zöngeciklus lefolyásával változik, az így kapott, a feszültségváltozást az idő függvényében rögzítő hullámforma, azaz az elektrolottogram a hangrés nyitódásáról és záródásáról, azaz a hangszalagok mozgásáról informál. Ugyanezen okból kifolyólag azonban az is igaz, hogy az EGG csak akkor igazán pontos, amikor a hangszalagok zárt állásban vannak (vagy éppen záródnak): a számos (az EGG használatát felderíteni szándékozó) exploratív vizsgálat (csak) azt bizonyította egyértelműen, hogy minél nagyobb a hangszalagok érintkezésének felülete, annál nagyobb az átmenő feszültség (HENRICH et al. 2004; HERBST et al. 2010). Az EGG-nek, valamint az EGG első deriváltjának, a dEGG-nek egy periódusa a zöngeciklus szakaszainak

²⁷ Az EGG rövidítés a szakirodalomban mind az elektrolottográfiát mint eljárást, mind pedig az így rögzített hullámformát, azaz az elektrolottogramot is jelöli.

megfelelően több elkülönülő részre osztható, ezt szemléltetjük a 28. ábrán (110 Hz alapfrekvencián, férfi ejtésében, modális és falzett zöngéképzésben, melyet a nemzetközi szakirodalomban bevett terminológia alapján az ábra forrását képező tanulmányban 1. és 2. fonációs mechanizmusnak neveznek, vö. még pl. TITZE 2004). Az ábrán az EGG és dEGG hullámformák mellett a hangrészen átáramló levegő mennyiségét az idő függvényében szemléltető áramlási glottogramot is láthatjuk, így egyrészt tanulmányozhatjuk, hogy hogyan függ össze a hangszalagok mozgását mutató EGG és dEGG az áramlási glottogrammal, másrészt megfigyelhetjük a hullámformák szinkronizálása szempontjából kitüntetett jelentőségű pontokat is. Az EGG hullámforma szakaszokra osztását HENRICH és munkatársainak (2004) összegző elemzése alapján ismertetjük.



28. ábra: Az EGG és dEGG a zöngéképzés egy ciklusában, és ezek összefüggése az áramlási glottogrammal a modális (1. mechanizmus) és falzett (2. mechanizmus) zöngéképzésben: (1)–(3): a hangszalagok összeháródása; (3)–(4): zárt állás; (4)–(6): a hangszalagok nyitódása; (6)–(1): nyitott állás (HENRICH et al. 2004: 1322 alapján)

A zöngeciklus szakaszainak megfelelően az ahhoz rendelt EGG egy ciklusa négy részre osztható az alábbiak szerint.

(1)–(3) a hangszalagok összezáródása: az EGG-n megfigyelhető feszültségnövekedés a hangszalagok közeledésére utal; a záródás a hangszalagok alsó éleinél kezdődik ((1)-től (2)-ig), majd továbbterjed a hangszalagok felső éléhez ((2)-től (3)-ig). Mivel a záródás általában gyorsabb, mint a nyitódás, ez a szakasz nagy (pozitív) meredekségű. A maximális meredekséget (*maximum slope*), azaz azt a pontot, amikor az elektromos feszültség a leggyorsabban változik a (2) pont jelöli. A későbbiekben még látni fogjuk, hogy ez a pont kitüntetett jelentőségű, így az elemzésben gyakran az ennek megállapítására kifejezetten alkalmas első deriváltat (dEGG) használjuk. (Mivel az első deriváltban a hirtelen változásokat a meredekségtől függően kisebb vagy nagyobb kiugrásokként látjuk, ez a pont a dEGG-ben egy nagyobb pozitív kiugrásként jelenik meg, lásd még HERBST et al. 2010.)²⁸

(3)–(4) zárt állás: a hangszalagok teljesen (vagy majdnem teljesen) összeérnek, így az átáramló feszültség maximális, míg a glottiszon átáramló levegő mennyisége nullára csökken (az itt tapasztalható csekély feszültségnövekedést vagy -csökkenést egyes kutatók az egymásnak csapódó rugalmas felületek dinamikájával magyarázzák).

(4)–(6) a hangszalagok nyitódása: az EGG-n megfigyelhető feszültségcsökkenés a hangszalagok távolodására utal; a hangszalagok alsó éleinek fokozatos szétválását ((4)-től (5)-ig) a felső élek szétválása követi ((5)-től (6)-ig). A maximális (negatív) meredekséget itt az (5)-ös pont jelöli, melyet (éppen ezért) a dEGG-n negatív csúcsként látunk. Mivel a hangszalagok szétválása lassabb, mint az összezáródás, ezért itt a maximális meredekség is kisebb, mint a ciklus másik felében, és a dEGG-n csak jóval kisebb csúcs jelenik meg, mint a záródási szakasz megfelelő pontján.

(6)–(1) nyitott állás: a hangszalagok nyitott állásban állnak. Az itt mért feszültség relatíve alacsony, és csak kismértékű ingadozást mutat (HENRICH et al. 2004).

Amint az a fenti leírásból is kiderül, a dEGG-n (2)-es számmal jelzett „záródási csúcs”, valamint az (5)-ös számmal jelzett „nyitódási csúcs” különös jelentőséggel bír. Ezekről ugyanis (párhuzamos videokimográfias és inverz szűrésen alapuló összehasonlító elemzéseken keresztül) kimutatták, hogy valóban a hangszalagok találkozási és szétválási pontjára utalnak (HENRICH et al. 2004). Mivel azonban a záródási szakasz gyorsabb, így a „záródási csúcs” valamivel egyértelműbb és megbízhatóbb referenciapontnak tekinthető. Ezt az információt fogjuk az ebben a fejezetben bemutatott akusztikai elemzésekben is hasznosítani.

²⁸ Megemlítendő, hogy bár a záródási pont prototipikusan egyetlen hirtelen kiugrás, egyes beszélőknél kétszúcsú is lehet. Ezt a kutatók más képalkotó eljárással készített összehasonlítások alapján általában a hangszalagok „cipzárszerű” záródása eredményének tartják, azaz a hangszalagok felületével párhuzamos előlről hátrafelé terjedő (*anterior-to-posterior*) vagy hátulról előrefelé terjedő (*posterior-to-anterior*) záródás lenyomatának (jöllehet részletes vizsgálatok híján a jelenségre valójában többféle magyarázatot is elképzelhetők tartanak; erről részletesebben lásd HENRICH et al. 2004; HERBST et al. 2010).

6.2. Az inverz szűrés, az áramlási glottogram, valamint az elektroglottográfia szerepe az inverz szűrésben

Az inverz szűrés a beszédképzés hagyományos, a forrást és szűrőt élesen elválasztó modelljéből indul ki, melynek alapfeltételezése a forrás és szűrő lineáris összefüggése is. Az eljárás elsődleges célja, hogy megbecsülje a glottison a beszédképzés (vagy más hangadás, illetve a légzés) alatt áthaladó levegő mennyiségét, valamint a levegő térfogatsebességét az idő függvényében. Ez a függvény az áramlási glottogram (*glottal flow, airflow*, vö. 28. ábra), mely tehát a hangrészen átáramló levegő mennyiségét ábrázolja miliméter/másodperc (mm/s) mennyiségként az idő függvényében. Az inverz szűrés során a toldalékcső átviteli karakterisztikáját leképező szűrősor inverzével, azaz az inverz szűrősorral kompenzáljuk a toldalékcső rezonátortevékenységet, így „állítva helyre” a gégénél létrejövő áramlási glottogramot. Az inverz szűrés lényege tehát a rezonanciák meghatározása és erősítő hatások kioltása az akusztikus jelben²⁹ (ROTHENBERG 1973). Éppen ezért, noha az eljárás elsődleges céljának nem ezt tekintjük, működési elvéből következően az inverz szűrés a rezonanciahelyek detektálására is használható.

Az inverz szűrést a jelen fejezetben bemutatott vizsgálatban is a formánsmeghatározás egyik módszereként alkalmazzuk, melyhez (HERTEGÁRD–GAUFFIN (1993) elemzéséhez hasonlóan) EGG-t is használunk. A módszertan megértéséhez a 28. ábrán az EGG és dEGG, valamint az áramlási glottogram összefüggéseit és megfeleltethetőségét kell alaposabban szemügyre vennünk. Ahogy már említettük, az EGG-ben kifejezett jelentőséggel bír a hangszalagok találkozására utaló pont, a dEGG-ben megfigyelhető „záródási csúcs”. A transzglottális levegőáram szempontjából ugyanis a záródási csúcs éppen annak a szakasznak a kezdetét jelöli, ahol a levegőáram mennyisége hirtelen nagyon lecsökken (ráadásul ez a pont az, amelyik az EGG- és dEGG-jelekben a zöngeciklus alatt a legegységelműbben azonosítható), tehát ez által a pont által mindkét hullámformán egyértelműen azonosítható a zárszakasz kezdete. Bár megfigyelhető az is, hogy a nyitódási csúccsal egy időben a levegőáram mennyisége is lassú növekedésnek indul, de jól látható, hogy az áramlási glottogram e szakaszának meredeksége jóval kisebb a záródási szakaszban tapasztalhatótnál. Ennek oka, hogy míg a záródás során a levegő útját hirtelen összezáródással elálló hangszalagok valóban képesek hirtelen csökkenést okozni a glottison átáramló levegő mennyiségében, addig az újra kinyíló hangrészen átáramló levegő sebessége a levegőoszlop nagyságától és tehetetlenségétől függ, így ezt a szakaszt jóval lassabban elinduló mozgás és kisebb áramlási sebesség jellemzi.

²⁹ Fontos itt megemlíteni, hogy a forrást és a szűrőt csak a modell értelmezésében, az elemzés megkönnyítése érdekében tekintjük függetlennek. Ezek valójában szoros kölcsönös függésben állnak egymással, amint arra a könyv korábbi pontjain már többször utaltunk. Ennek értelmében tehát az inverz szűrés során visszafejtett jel is csak elméleti reprezentációja lehet a hangrészen átáramló levegőmennyiség időbeli vetületének. Igaz ugyanakkor az is, hogy sem a hangrészen átáramló levegő, sem pedig a gégénél keletkező hangnyomásingadozás, azaz a zöngé nem mérhető vagy rögzíthető közvetlenül, ilyen értelemben tehát ezekről a legjobb közelítést az inverz szűrés által kaphatjuk.

Az inverz szűrés lényege, hogy a toldalékcso rezonátortevékenységét kompenzálva helyreállítsuk a gégénéél létrejövő áramlási glottogram hullámformáját, ami elsősorban a fodrozódásmentes zárszakasz-hullámforma helyreállítását jelenti. Ebben pedig segítség lehet a dEGG, hiszen az a záródási csúcson jelzi azt a pontot, ahol a zárszakasz kezdődik, és ahová az áramlási glottogram a zárszakasz kezdetét jelző hajlatának esnie kell (ez lesz tehát a (2)-es számmal jelölt pont). Az EGG és különösen a dEGG tehát nagy segítség a manuális inverz szűrés során, különösen akkor, ha magasabb alapfrekvenciákon ejtett hangokat elemzünk falzett regiszterben (vagy más szóval a zöngképzés 2. mechanizmusában), melyek áramlási glottogramjának hullámformája a modális, alacsony alapfrekvenciájú zöngé hullámformájánál rövidebb hullámhosszal és rövidebb zárszakasszal jellemezhető (lásd pl. HENRICH et al. 2004, valamint a 28. ábrát).

Az inverz szűrést a különböző alapfrekvenciájú énekelt magánhangzók formánshelyeinek meghatározására eddig HERTEGÁRD és GAUFFIN (1993), valamint SUNDBERG és munkatársai (2011, 2013) használták sikerrel. Míg az eljárást HERTEGÁRD és GAUFFIN (1993) egy női (szoprán) énekes magas alapfrekvenciájú magánhangzóinak mérésében alkalmazta, SUNDBERG és munkatársai (2011, 2013) férfi hangosztályok alacsonyabb alapfrekvenciájú énekprodukciójában vetették be. Az esetünkben leginkább relevánsnak tekinthető, HERTEGÁRD és GAUFFIN (1993) által végzett vizsgálat fő konklúziója az, hogy az inverz szűrés EGG alkalmazásával megfelelő módszer lehet még a magas alapfrekvencián ejtett magánhangzók formánsainak detektálásában is.

Hertegárd és Gauffin (1993) nyomán a jelen fejezetben bemutatott kísérletben a magas alapfrekvenciájú szoprán énekülésben megvalósult magánhangzók formánshelyeiket inverz szűréssel állapítjuk meg, és ehhez párhuzamosan rögzített EGG-jelet is használunk. Ugyanakkor, mivel az áramlási glottogram hullámformája a magasabb alapfrekvenciákon jelentősen megváltozik, még ez az eljárás is tartogat némi bizonytalanságot, ezért ezeket az inverz szűréssel mért adatokat validáljuk, illetve a mérési módszert kiegészítjük a 6.3. alfejezetben részletesebben is ismertetett módon, egy külső gerjesztésű jel használatával.

6.3. A kísérlet anyaga és módszere, valamint a kísérleti személyek

6.3.1. Felvételkedészítés, az akusztikai vizsgálat anyaga

A kísérletben három professzionális szinten éneklő, praktizáló szoprán szólóénekesnő (S1, S2, S3, lásd a 7. táblázatot) énekprodukcióját vizsgáltuk a kilenc hosszú vagy hosszán ejthető magyar köznyelvi magánhangzó /v a: ε e: i: o: ø: u: y:/ ejtése közben, hat énekelt alap-

frekvencián (három oktáv hangterjedelemben) 175 és 988 Hz között (az *f, h, f', h', f'', h''* zenei hangokon, lásd a 6. táblázatban) és beszéd módban.

7. táblázat: A kísérletben részt vevő énekesek

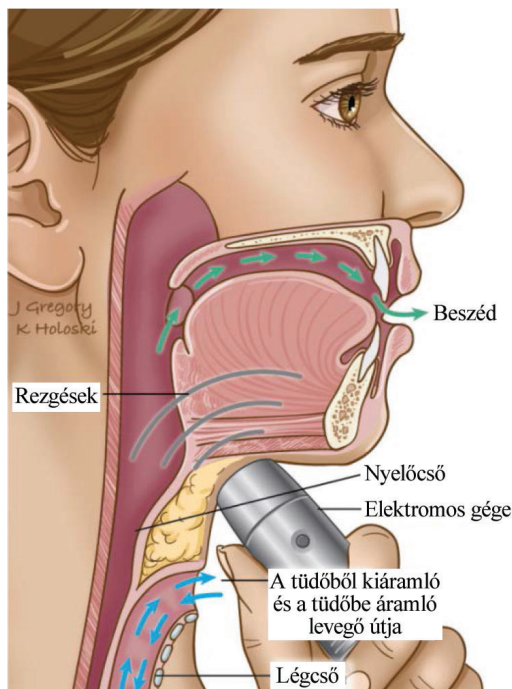
Az énekes kódja	Életkor	Hol tanult?	Milyen formában, hol énekel?	Mióta énekel?
S1	28	Liszt Ferenc Zeneművészeti Egyetem	Szólistaként, kórusban, énektanárként is dolgozik	14 éve
S2	45	Liszt Ferenc Zeneművészeti Egyetem	Szólistaként, kórusban, énektanárként is dolgozik	Több mint 30 éve
S3	28	Magánénektanárnál	Szólistaként, kórusban	14 éve

Ahogy eddig is, mivel a vizsgálatban a paraméterek szigorú kontrollálása, valamint a felvételi és elemzési protokoll sajátosságai miatt csak kitartott hangokat vizsgáltunk, öt fonológiai hosszúsági pár rövid tagja (az /i o ø u y/ magánhangzók) kimaradt az elemzésből. Ennek ellenére úgy tekinthetjük, hogy a jelen vizsgálati anyag reprezentálja a teljes magyar köznyelvi magánhangzókészletet, hiszen az itt kimaradt magánhangzókat hagyományosan csak időtartamukban tartjuk a pár hosszú tagjától eltérőnek a magyarban (lásd 1.1.4.4. fejezet). Az énekesek minden magánhangzót minden alapfrekvencián kétszer ejtettek ki, így összesen (3 énekes × 9 magánhangzó × 7 alapfrekvencia × 2 ismétlés ⇒) 378 stimulust vizsgáltunk. Minden alapfrekvencián, minden kilencelemű magánhangzócsoport felvétele előtt az énekesek referenciahangot hallottak fülhallgatón keresztül, így biztosítottuk a pontos intonációt.

A vizsgálat módszereiben az énekelt magánhangzók formánsainak meghatározásához két, korábban már sikerrel használt elemzési technikát ötvöztünk. Az első a magas alapfrekvenciájú zöngé helyettesítése és a toldalékcső „újra-mintavételezése” volt egy mély alapfrekvenciájú, nagy felhangsűrűségű külső gerjesztésű jel segítségével. A másik pedig a beszédjel (azaz az énekesek hangjának) inverz szűrése volt, melyhez EGG-t, illetve dEGG-t is használtunk. Ennek megfelelően a felvételek során két jelet rögzítettünk párhuzamosan, két csatornán. Az első csatornán az énekes száját elhagyó hangot rögzítettük (audió) egy gömbkarakterisztikájú mikrofonnal, melyet 30 cm távolságra helyeztünk el az énekes szájától. (A mikrofont a szájníllástól egy bukósisakra szerelt pálcá tartotta állandó távolságra, melyet az énekesek a felvétel során viseltek.) A másik csatornán pedig a hangszalagok mozgását rögzítettük egy elektrolottográf³⁰ segítségével.

³⁰ A vizsgálatban egy *Laryngograph* márkájú elektrolottográfot és az ahhoz tartozó *Speech Studio* szoftvert használtam (Laryngograph, Laryngograph® Ltd, Wallington, Greater London, UK). Mind az elektrolottográfot, mind pedig a felvételek helyszínénél szolgáló csendesített szobát az MTA Nyelvtudományi Intézetének Fonetikai osztályán működő Kempelen Farkas Beszédkutató Laboratórium bocsátotta rendelkezésemre, amit ezúton is hálásan köszönök, illetve köszönöm Gráczy Tekla Etelkának és Beke Andrásnak a felvételkészítésben nyújtott segítségüket.

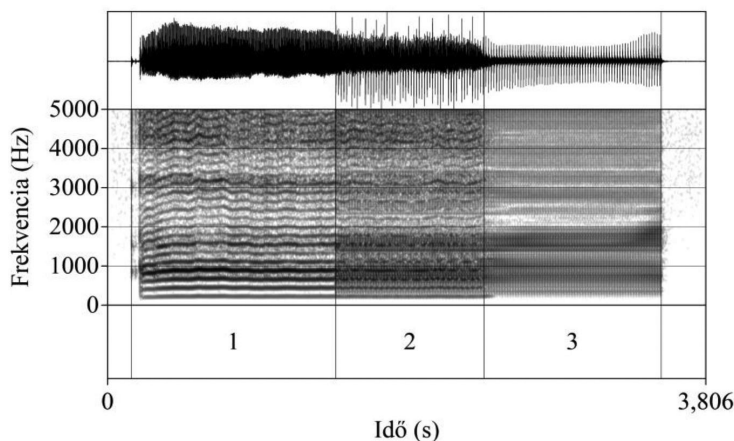
Ahhoz, hogy a toldalékcsovön áthaladó, a száját elhagyó kimeneti jelben is mérhető rezonanciákat kapjunk, az egyes hangok ejtésére jellemző artikulációs konfigurációt az énekes saját fonációja mellett egy elektromos gégeből (vagy más néven *elektrolarynx*ből) származó külső gerjesztésű, mély alulfrekvenciájú, nagy felhangsűrűségű jellel is „mintavételeztük”. Az elektromos géget szokásosan a hagyományos módon valamilyen okból (általában részleges vagy egészes gégeeltávolítás miatt) zöngét képezni nem tudó személyek használják. Az eszköz mély, berregő hangot bocsát ki, melyet használója a gégeszint felett a nyak szöveteihez szorítva a toldalékcsőbe juttathat, a rezgés pedig (a zöngéhez hasonlóan) a garatüregen át a szájnyílás felé terjed tovább, mígnem végül távozik a szájnyíláson keresztül. A száját elhagyó, az elektromos gége gerjesztette hang spektrumát a toldalékcső üregei – a beszédhangképzéssel analóg módon – rezonátortevékenységük révén módosítják, így a kimenetben beszédhangszerű hangok észlelhetők és rögzíthetők (lásd 29. ábra). Az így keletkező akusztikai jel elemzése tehát alkalmas arra, hogy benne az artikulációs konfigurációra az ejtés közben jellemző formánsfrekvenciákat meghatározzuk.



29. ábra: Az elektromos gége működési elve gégeeltávolításon átesett páciensről készült keresztmetszeti ábrán szemléltetve. Az elektromos gége keltette berregő hang a nyak szövetein keresztül a toldalékcsőbe jut, majd a garatüregen és a szájüregen is áthaladva távozik a szájnyíláson keresztül. A hang spektrumát a toldalékcső üregei rezonátortevékenységük révén módosítják, így a kimenetben beszédhangszerű hangok észlelhetők és rögzíthetők (az ábra a <http://www.headandneckcancerguide.org/> alapján készült)

Az elektromos gégével történt felvételekhez az énekeseket kitartott hangok éneklésére kértük. A felvétel folyamatát és szakaszait a 30. ábrán szemléltetjük, és ez alapján is mutatjuk be. Az ábra egy beszédben ejtett /a:/ magánhangzó spektrogramját mutatja a kétféle mintavételezési eljárással készült felvételen.

Az énekesek feladata az volt, hogy a magánhangzó ejtése közben néhány másodpercig énekeljenek természetesen, azaz képezzenek zöngét (30. ábra 1-es szakasz), majd kapcsolják be az ekkor már a nyakukhoz szorított elektromos gégét (30. ábra 2-es szakasz), végül pedig merevítsék meg az artikulációjukat, és hirtelen hagyják abba a zöngéképzést úgy, hogy a hangrést a levegő visszatartásával bezárják (30. ábra 3-as szakasz).



30. ábra: Az /a:/ hangzó beszédbeli ejtéséről készült keskeny sávú spektrogram S1 énekes ejtésében.

A spektrogram a felvételek során használt kétféle „mintavételezési” eljárást (valamint azok átlapolását) szemlélteti. (1) az énekes zöngét képez; (2) az énekes zöngéképzés közben bekapcsolja a gégeszint felett a nyakához illesztett elektromos gégét; (3) az énekes hirtelen abbahagyja a zöngéképzést, visszatartja a levegőt, bezárja a hangrést, így a felvétel már csak az elektromos gégeből származó, az énekes száján keresztül távozó hanghullámokat rögzíti

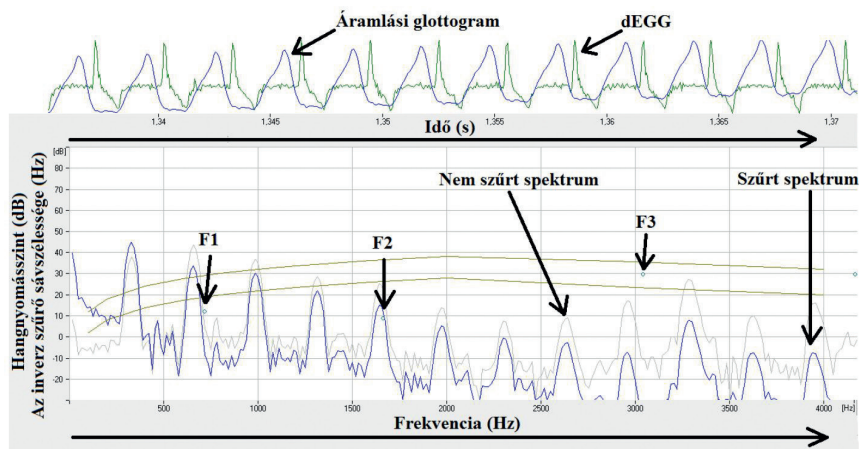
A (2)-es számmal jelölt szakaszon tehát az énekes toldaléksövét mind saját zöngéje, mind pedig az elektromos gége hangja „mintavételezte”, míg a (3)-as számmal jelölt szakaszon már csak az elektromos gégének a toldalékső által módosított hanghullámai hagyták el a szájüreget. Mivel a hangrést az énekesek ez utóbbi, azaz a (3)-as szakaszban bezárták, így a szubglottális területek rezonátortevékenysége itt nem lehetett hatással a kimenő jel spektrumára. A bukósisakra rögzített mikrofon segítségével tehát a 30. ábrán látható módon ejtett magánhangzókat vettünk fel. Az elektromos gége használatát a felvételek elkészítése előtt a szükséges ideig minden énekessel gyakoroltuk.

6.3.2. Akusztikai elemzések

A kísérletben a természetes módon (zöngéképzés közben) produkált (az ének és beszéd módban) rögzített magánhangzószeleteket (lásd 30. *ábra* 1-es szakasz) inverz szűrés segítségével elemeztük a DeCap szoftverben (GRANQVIST é. n.).³¹ A szűréshez a beszédjelet és az EGG-jelet tartalmazó hangsávokat először időben illesztettük. Mivel a két jelfelvételi pont, azaz az EGG-elektrodák (vagyis a hangszalagok) és a mikrofon közti távolság nem nulla, ezért a két sávot a felvételkészítéskor szükségszerűen egymáshoz képest időbeli elcsúszásban rögzítettük. Ezt a (közel 45 cm-es) távolságot a jelek szinkronizálásához úgy kompenzáltuk, hogy kiszámítottuk, hogy a hang terjedési sebességéből, valamint a két pont közti távolságból következően mekkora idői eltolódásban rögzítettük a jelet a két csatornán, és a kapott értékkel eltoltuk (azaz „előrébb hoztuk”) az EGG-csatornát. (Mivel a toldalékcső felnőtt nőknél átlagosan 15 cm, és ehhez adódott hozzá a mikrofon távolsága az ajkaktól, így 45 cm távolsággal számoltunk.) Az eltolás mértékét matematikai úton ($t = \Delta s / v = 45 \text{ cm} / 35\,000 \text{ cm/s} = 1,3 \text{ ms}$ -ban állapítottuk meg, amit az első ellenőrző mérések (ti. a beszédjelek inverz szűrése) után 2 ms-ra korrigáltunk (úgy, hogy a dEGG-ben detektálható záródási csúcs, valamint az áramlási glottogram zárszakaszának kezdete tökéletes szinkronba kerüljön).

A DeCap szoftver manuális inverz szűrést tesz lehetővé. A program egyszerre kezeli az EGG-t és a mikrofon által felvett audiojelet. Az EGG-t (és deriváltját, a dEGG-t) a felület felső ablakában hullámformaként jeleníti meg, a mikrofon által rögzített jelet pedig egyszerre hullámformaként (szintén a felső ablakban) és spektrális felbontásban (az alsó ablakban) is látjuk (31. *ábra*). Az elemzéshez a program integrálja a mikrofonnal rögzített hangnyomás-idő függvényt, azaz a hangnyomás-idő függvényt levegőáramlás-idő függvényé alakítja (abból kiindulva, hogy a levegőáramlás-idő függvény deriváltja a hangnyomás-idő függvény). Ezt a jelet szűrjük az elemzés során. Emellett pedig az EGG is deriválható (tetszés szerint), tehát a dEGG jel is megjeleníthető. Az elemzésben a 31. *ábrán* F_1 , F_2 , F_3 jelöléssel jelzett formánsokat, valamint (szükség szerint) a további, magasabb frekvenciájú és sorszámú formánsokat manuálisan hangoljuk: vízszintesen állítjuk a formáns frekvenciáját, függőlegesen pedig a formáns sáv szélességét. A formánsok (azaz a szűrők) manuális hangolásának megfelelően mind a hullámformát, mind pedig a spektrumot valós időben módosítja a program (a klasszikus egyenleteknek megfelelően, lásd FANT 1960), tehát a szűrés eredménye azonnal látható lesz, miközben akár a nem szűrt, eredeti jelet is megjeleníthetjük (az *ábrán* ez látható halványabb színnel az alsó, FFT-ablakban).

³¹ Az elemzéseket a KTH Royal Institute of Technology, Speech Music and Hearing tanszékén végeztem prof. emer. Johan Sundberg és dr. Svante Granqvist segítségével egy egy hónapos ösztöndíjas tanulmányút keretében. A segítséget ezúton is hálásan köszönöm a kutatóknak, a Campus Hungary programnak és a Balassi Intézetnek pedig a támogatást.



31. ábra: Az inverz szűrés menete a DeCap programban. A felső ablakban a dEGG (zöld vonal), valamint a szűrt áramlási glottogram (a hangnyomás-idő függvény integráltja) (kék vonal) látható. Az alsó ablakban az audiófelvétel eredeti (szürke vonal) és szűrt (kék vonal) spektrumát, valamint a beállított szűrők (apró körök; F_1 , F_2 , F_3) frekvenciáját és sávzélességét figyelhetjük (SUNDBERG et al. 2013 alapján)

Az inverz szűrés, azaz a szűrők manuális hangolása során a cél 1. a dEGG-ben azonosítható záródási csúcs és az áramlási glottogramon beazonosítható zárszakasz kezdetének³² szinkronizálása, 2. a zárszakasz-hullámforma fodrozódásainak megszüntetése, és 3. a spektrum meredekségének kiegyenlítése a lokális maximumok és minimumok lehetőség szerinti teljes kisimításával (vö. HERTEGÁRD–GAUFFIN 1993; SUNDBERG 2011, 2013). A szűrők helyes beállításával az áramlási glottogram a glottison áthaladó levegő (elméleti) hullámformáját rajzolja ki (vö. 28. ábra). A 31. ábrán látható áramlási glottogram hullámformáján (kék vonal) a szűrés utáni állapotot láthatjuk egy jól beállított szűrősor alkalmazása esetén. Az alsó, FFT-ablakban látható közel párhuzamos és megközelítőleg vízszintes vonalak (a „kerítés”) az elemzést segítő az adott frekvenciákon reális formánsávzsélességeket jelölik. Ezekről, női beszélő lévén gyakran eltértek a kapott adatok, a legtöbb esetben ugyanis a javasoltnál nagyobb sávzsélességeket mértünk az első két formáns esetében (azaz a formások értékét mutató jelölők a „kerítés” alatt helyezkedtek el, amint azt a 31. ábrán is megfigyelhetjük).

Az elektromos gégével rögzített hanganyagokat (azaz a 30. ábrán a 3-as számmal jelölt hangzórészleteket) a DeCap (GRANQVIST é. n.) és a Praat szoftver (BOERSMA–WEENINK 2009) segítségével elemeztük. Az itt kapott formánsértékeket összevetettük az inverz szűréssel kapott formánsértékekkel. A kétféle módszerrel nyert adatok eltérése esetén ellenőriztük a lehetséges okokat, és javítottuk a hibákat, egyezés esetén pedig elfogadtuk a mérést – a két

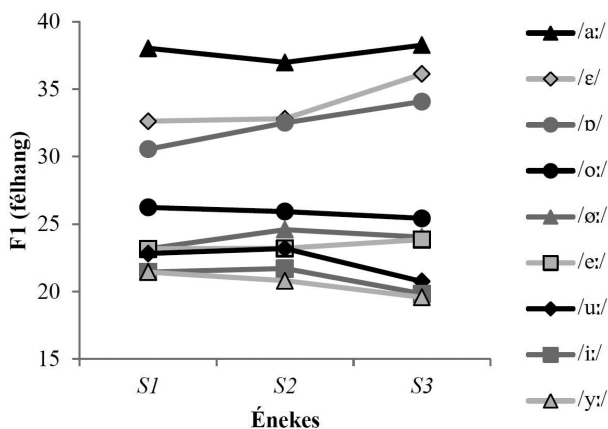
³² SUNDBERG (2011, 2013) erre a pontra az MFDR megnevezéssel (*Maximum Flow Declination Rate*) hivatkozik, mely kifejezés az áramlási glottogram maximális meredekségére utal a záródási (azaz deklinációs) szakaszban.

mérés mód tehát egymás ellenőrzéseként szolgált az elemzésben. Mivel tudjuk azt, hogy a hallási észlelési folyamatok frekvenciafüggőek, az akusztikai mérések eredményeképpen kapott adatokat érzetileg (félhangokra) transzformáltuk, minden esetben 100 Hz frekvenciához viszonyítva (ezekre hivatkozunk majd a 7. fejezet észlelési kísérletének elemézéskor is).

6.4. Eredmények

Mivel az S2 kódjelű énekes a felvételek során (fáradtság miatt) nem tudta megfelelő hangszínezettel végigénekelni a stimulusokat a legmagasabb, azaz a *h''* (988 Hz) zenei alaphangfrekvencián, ezért az eredményeket szemléltető ábrákon az S2 kódjelű adatsorokról a megfelelő adatok hiányoznak. Ugyanezen okból a 34. ábrán látható magánhangzótér-változásokat is a legmagasabbnál egyvel alacsonyabb (*f''*, azaz 698 Hz) alaphangfrekvencián szemléltetjük, ahol mind a három énekes összes adata rendelkezésre állt.

A 32. ábra a beszéd módban ejtett magánhangzó-megvalósulások első formánsának értékét mutatja (félhangokban). Látható, hogy a három énekesnek a beszéd módban ejtett magánhangzói hasonló F_1 -értékkel realizálódtak. A beszélők (énekesek) közti kisebb-nagyobb eltérések ellenére a nyíltabb hangzók között nagyjából egyértelműen felállítható sorrendet látunk: a legnyíltabban, legmagasabb F_1 -értékkel az /a:/ hangot ejtették az énekesek, melyet rendre az /ε/ és /ɒ/ követett (a nyíltságot természetesen itt szigorúan akusztikailag értve). A zárttabban ejtett hangzók közti különbségek már nem ilyen egyértelműek, de az minden énekes esetében megfigyelhető, hogy az alsó nyelvállásúaknál valamivel alacsonyabb F_1 -értékkel a középső nyelvállású magánhangzók realizálódtak (az /o:/, /ø:/ és /e:/), míg a legalacsonyabb F_1 -értékkel a felső nyelvállású magánhangzók (az /u:/, /i:/ és /y:/).

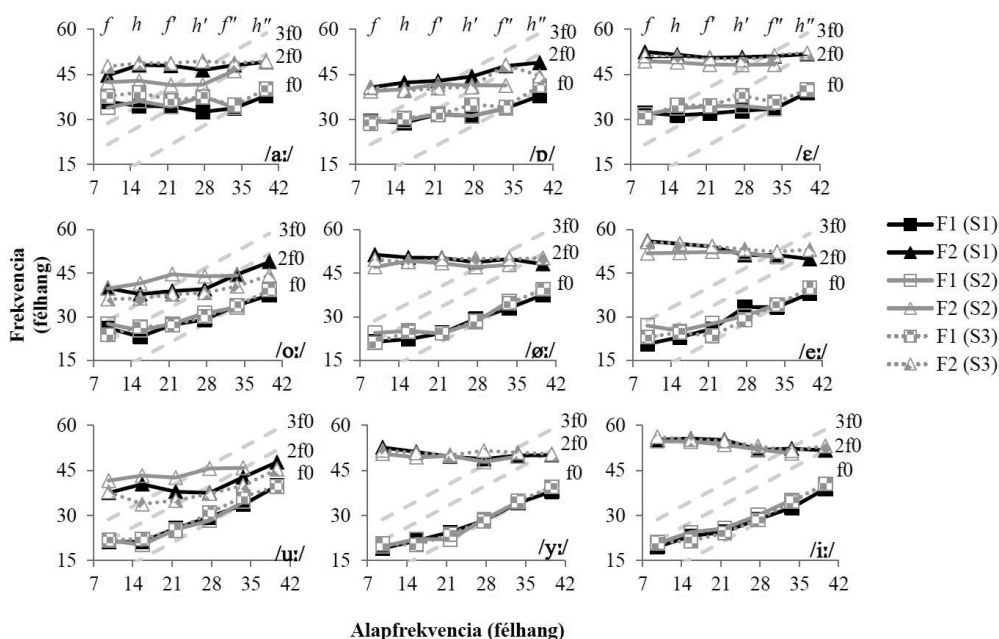


32. ábra: A beszédben ejtett magánhangzók első formánsa a három énekes ejtésében

6.4.1. Az énekelt magánhangzók első formánsának alakulása

A 33. ábra a kilenc vizsgált magánhangzó első és második formánsának értékeit mutatja (y tengely) az alapfrekvencia függvényében (x tengely), félhangokban mérve, a három énekes ejtésében külön-külön. Az ábrákon világos színű szaggatott vonallal jelöltük az alapfrekvenciát és annak egész számú többszöröseit, azaz a felhangokat vagy felharmonikusokat ($f_0, 2f_0, 3f_0$ stb.). Az énekesek közti kisebb mértékű variabilitás ellenére egyértelmű tendenciákat találtunk: mindhárom énekes ejtésében megjelent az F_1 az alapfrekvenciára (vagy ahhoz közel eső frekvenciaértékre) hangolása akkor, ha az f_0 megközelítette az F_1 -nek a beszédben az adott magánhangzóra jellemző értékét (vö. 32. ábra). Attól a ponttól kezdve, ahol az f_0 az (alapfrekvencia-emelés során) elérte az F_1 -nek a beszédben jellemző régióját, az $F_1 : f_0$ hangolást minden énekes alkalmazta minden magasabb f_0 és minden magánhangzó esetében. A hangolási tendencia megjelenésének alsó határa tehát az F_1 -től, azaz az adott hangzó nyíltságtól függött. Ez az érték a legmagasabb (h'' , 988 Hz) a legnyitottabb /a: ɒ ε/ esetében volt, ennél alacsonyabban tapasztaltuk (h' , 494 Hz) a második nyíltsági fokú /o: ø: e:/ esetében, míg a legalacsonyabban (f' , 349 Hz) az első nyíltsági fokú /u: y: i:/ esetében láttuk megjelenni. A hangolási tendencia megjelenése előtt (tehát a fenti alapfrekvenciáknál alacsonyabb frekvenciákon) az F_1 frekvenciaértéke az f_0 -tól függetlenebbnek mutatkozott, és a beszédhez közeli frekvenciaregióban realizálódott.

Az $F_1 : f_0$ hangolási stratégia felső határa tekintetében az eredményeink nem egyeznek JOLIVEAU és munkatársainak (2004) eredményeivel. Az idézett szerzők azt találták, hogy az $F_1 : f_0$ hangolásnak van felső határa, mégpedig az /ɔ u/ esetében valahol 800 és 1000 Hz között: ezeken a frekvenciákon (és ezek fölött) vizsgálatukban az ajakkerekítéses /ɔ u/ hangzók első formánsa már az f_0 értéke alatt volt mérhető. A jelen vizsgálatban azonban nem tapasztaltuk azt, hogy az énekesek elérték volna az $F_1 : f_0$ stratégia alkalmazásának felső határát, sem az ajakréses, sem az ajakkerekítéses hangoknál: a zenei f'' hang (988 Hz alapfrekvencia) elérésével mindhárom, a kísérletben részt vevő énekesnél továbbra is az f_0 értékének közelében mértük az F_1 értékét.



33. ábra: A három szoprán énekes által ejtett beszédhangok F_1 és F_2 formánsfrekvenciái (y tengely) az alapfrekvencia (x tengely) függvényében. Minden jelölő egy énekes egy adott magánhangzójának egy adott alapfrekvencián ejtett megvalósulását mutatja.

Világosszürke szaggatott vonalakkal az alapfrekvenciát és annak első két felharmonikusát jelöltük (rendre f_0 , $2f_0$, $3f_0$)

6.4.2. Az énekelt magánhangzók második formánsának alakulása

Az F_2 értéke az F_1 -nél jóval függetlenebbnek mutatkozott az alapfrekvenciától, különösen az elől képzett /ɛ/, valamint a centrális képzésű /a:/ hangzók esetében. Az elől képzett /e: i: y:/ F_2 -értékében az $F_1 : f_0$ hangolásának megjelenésével kezdődően az f_0 (és F_1) emelésével párhuzamos, fokozatos, de enyhe csökkenést tapasztaltunk az alapfrekvencia emelésével. A hátul képzett /ɒ o: u:/ F_2 -je azonban (az elől képzett magánhangzókhoz képest) szorosabb összefüggést mutatott az alapfrekvenciával, ugyanis e hangok F_2 -értéke (az $F_1 : f_0$ hangolás megjelenésének alapfrekvenciájától fölfelé, az f_0 emelésével párhuzamos) fokozatos és nagymértékű növekedést mutatott.

Az F_2 -értéknek változására jellemző tendenciák az akusztikai magánhangzótér kihasználtsága szempontjából azt jelentik, hogy az elől és hátul képzett magánhangzók az alapfrekvencia emelésével az $F_1 : f_0$ hangolás megjelenésétől, tehát nagyjából az f' és h' zenei hangoktól (körülbelül 300-500 Hz alapfrekvenciától) kezdve egyre kevésbé különülnek el egymástól,

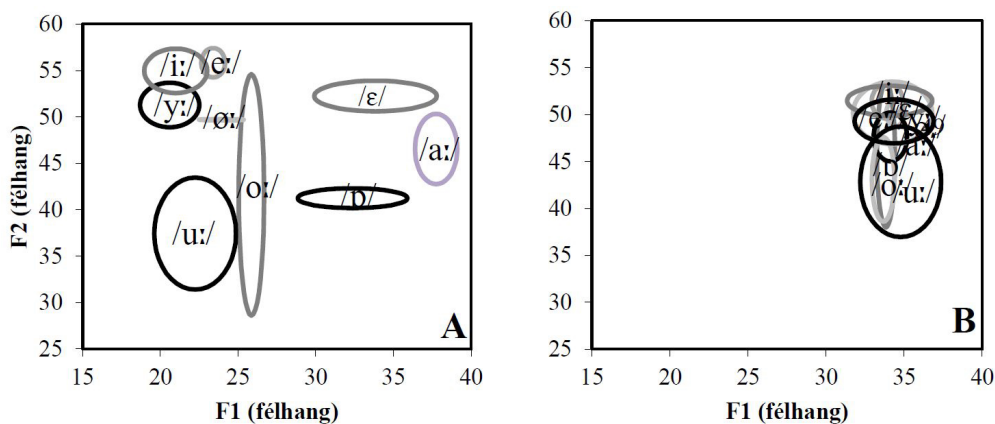
a konvergálást pedig elsősorban a veláris hangok fokozatosan változó, egyre palatálisabb képzése okozza. Az F_2 változása tehát a vízszintes nyelv helyzet függvényében eltérő módon volt jellemző a magánhangzókra, de az F_2 változásában (azaz az elől-hátul képzettség változásában) a nyíltsági foknak is szerepe van: a palatálisok között az első és második nyíltsági fokú /e: i:/, míg a velárisok között az ugyanezen nyíltsági fokokba tartozó /o: u:/ F_2 -értéke függött a leginkább az f_0 -tól, és változott a legjelentősebben az alaphangfrekvencia emelésével.

6.4.3. Az akusztikai magánhangzótér az alaphangfrekvencia függvényében

Az akusztikai magánhangzótérnek az előzőekben már részben érintett, magas alaphangfrekvencián tapasztalt változásait a 34. ábrán szemléltetjük. Az ábrán szereplő koordináta-rendszerek két tengelyét az F_1 és F_2 magánhangzóformánsok adják (ezeket rendre az x és az y tengelyen ábrázoltuk), így tehát az x tengely a nyíltsági fokkal (azaz közvetve a függőleges nyelv helyzettel), az y tengely pedig a veláris és palatális képzéshellyel, azaz a vízszintes nyelv helyzettel függ össze. Ilyen módon az x tengely mentén balról jobbra haladva találjuk rendre az első, második, harmadik és negyedik nyíltsági fokú magánhangzókat, míg az y tengelyen fentről lefelé haladva a veláris és palatális magánhangzók különülnek el egymástól. Az A panel a beszéd módban ejtett magánhangzók eloszlását, a B panel pedig a 698 Hz alaphangfrekvencián, azaz a zenei f'' hangon énekelt magánhangzók megvalósulásait mutatja az akusztikai magánhangzótérben a három énekes ejtésében: az ábrákon az egyes magánhangzó-megvalósulásokat reprezentáló ellipszisek középpontjainak koordinátáit az előzőekben ismertetett F_1 - és F_2 -értékek átlagával, míg az ellipszisek sugarát a megfelelő adatok kétszeres szórásával határoztuk meg.

Az A panelen látható, hogy a beszédben megvalósult magánhangzók az érzetileg transzformált (félhangokat, azaz logaritmikus skálát alkalmazó) pszichoakusztikai térben jól elkülönülnek egymástól: az ellipszisek, tehát a beszédhangok megvalósulásai közti átfedések csak kismértékben jellemzők. A B panelen ezzel szemben a magánhangzótér zsugorodását és a nyílt és centrális ejtés felé, azaz az /a:/ felé történt elmozdulását figyelhetjük meg az f'' zenei hangon (ezt az ábrán az ellipszisek horizontális irányban jobbra, valamint vertikális irányban középtájra történt elmozdulásából olvashatjuk le).

Bár az f'' zenei hang nem számít kifejezetten magasnak a szopránok hangtartományát tekintve, már ezen az alaphangfrekvencián is tapasztalható a veláris és palatális hangok akusztikai elkülönítésének drasztikus csökkenése, és a hangzók akusztikai nyitódása, amelyet az alaphangfrekvencia extrém emelésével vártunk. Ugyanakkor az is látható, hogy az elől és hátul képzett hangok ezen a közepesen magas alaphangfrekvencián még valamivel nagyobb mértékben elkülönülnek egymástól (mint például a különböző nyíltsági fokú hangzók).



34. ábra: A három énekes által énekelt magánhangzók megvalósulásának eloszlása a (pszicho)akusztikai magánhangzótérben a beszédben (A panel) és az f'' zenei alaphangon, 698 Hz alapfrekvencián (B panel). A magánhangzók megvalósulásait reprezentáló ellipszisek középpontját az F_1 - és F_2 -értékek átlagával, az ellipszisek sugarát pedig a megfelelő adatok kétszeres szórásával határoztuk meg

6.5. Következtetések

A fejezetben közölt vizsgálatban professzionális szinten éneklő, praktizáló szoprán énekeseknek a magánhangzókat érintő formánshangolási stratégiáit vizsgáltuk magas alapfrekvenciájú éneklésben. Megelőző akusztikai vizsgálatunkkal (2. fejezet) ellentétben ez alkalommal korábban leírt módszerek egy új kombinációjának segítségével módunk nyílt nem csak az elhangzó beszédhangok spektrális maximumhelyeinek detekciójára, de a produkció közben a rezonátorüreg-rendszerre jellemző sajátfrekvenciák, azaz a formánsok megállapítására is. Ezzel a fejezetben ismertetett vizsgálat tekinthető az első olyan akusztikai elemzésnek, mely egy adott nyelv összes magánhangzójának az alapfrekvencia emelésével megváltozó produkcióját feltérképezte (közvetlen, akusztikai adatokon keresztül, vö. 1.2. alfejezet), különös tekintettel az $F_1 : f_0$ hangolásra.

Magas alapfrekvencián a szoprán operaénekesek gyakran annál a frekvenciasávnál is magasabb alapfrekvenciájú zenei hangok ejtésére kényszerülnek, amelyen a beszélt ejtésben produkált magánhangzók hangszínét jelentősen meghatározó első spektrális maximum, az első formáns realizálódik. Ha viszont az f_0 értéke ilyen módon magasabbra kerül, mint az F_1 -é, akkor a szájat elhagyó akusztikai jel frekvencia-összetevőinek energiája, tehát az ejtett hang

hangossága is jelentősen lecsökken, a hangzó hangszínezete pedig hirtelen és drasztikus (az éneklésben nem kívánatos) változást szenved el. Bár a hangosságélményt érintő nem kívánt hatás megnövelt szubglottális nyomással némileg kompenzálható lenne – azaz hangerő-növekedést idézne elő, ha az énekes a fent említett helyzetben több levegőt és nagyobb intenzitással préselne ki a tüdejéből a hangadáshoz –, a nagyobb szubglottális nyomással fenntartott, huzamosabb fonáció a hangszalagok megerőltetését okozná, tehát hosszú távon káros, és a jó technikájú éneklésben mindenképpen került megoldás. Éppen ezért az éneklés szakirodalmában elterjedt nézet szerint az énekes e helyett a káros technika helyett a megemelt alapprofrekvencia hatására fellépő hangerőcsökkenést és hangszínváltozást a formánsok frekvenciáit befolyásoló stratégiákkal kompenzálja. Ez pedig elsősorban az $F_1 : f_0$ hangolás, melynek megjelenésére ugyanakkor (amint azt az 1.2. alfejezetben részletesen tárgyaltuk) a mérőmódszertan korlátai miatt viszonylag kevés empirikus bizonyítékkal rendelkezünk. Ráadásul a korábbi mérések egyike sem célozta meg egy adott nyelv teljes magánhangzókészletének vizsgálatát, így a beszédhangok az f_0 -tól függő változásának rendszer szinten érvényesülő tendenciáiról eddig meglehetősen keveset tudtunk.

Az itt bemutatott vizsgálat eredményei alapján megfogalmazhatjuk, hogy a kísérletben vizsgált három profi szoprán énekes egységes és általánosan leírható formánsstratégiákat alkalmazott, továbbá az eredmények a korábban, más metodológiával és más nyelvek hangzóira kapott eredményekkel is összhangban vannak. A jelen vizsgálatban igen nagy mennyiségű anyagon, a magyar nyelv összes magánhangzóján, széles (a szopránok hangterjedelmét lefedő, három oktávnyi) hangterjedelemben, három énekes produkciójában sikerült demonstrálnunk azt a nemzetközi szakirodalomban még csak néhány nyelv bizonyos hangzóira leírt stratégiát, mely szerint azokban az esetekben, amikor az f_0 megközelíti az adott magánhangzó beszédbeli ejtésére jellemző F_1 -értékét, az énekes az F_1 formánst a megemelt f_0 értékére (vagy valamivel a fölé) emeli.

Fontosnak tartjuk hangsúlyozni, hogy bár az adataink az F_1 formáns hangolását egyértelműen alátámasztják, a formáns frekvenciaértékének megváltoztatásához használt artikulációs stratégiákat pusztán a jelen vizsgálat akusztikai adataiból nem következtethetjük ki. Adataink tehát csak és kizárólag a rezonanciák hangolásának akusztikai alapú stratégiáit demonstrálják, nem pedig azokat az artikulációs stratégiákat, melyeket az énekesek ezen akusztikai célok eléréséhez alkalmaznak. Ennek megfelelően tehát abban az éneklés szakirodalmában vitatottnak tűnő kérdésben sem tudunk állást foglalni, mely az F_1 hangolását előidéző artikulátorok mibenlétét firtatja a különböző nyelvállásfokú hangok esetében: nem rendelkezünk ugyanis információval arra nézvést, hogy a kérdéses akusztikai stratégiát az állkapocsnyitás szögének növelése vagy inkább a nyelvállásfok csökkentése okozza a zárt ejtésű, felső nyelvállású magánhangzóknál (vö. SUNDBERG–SKOOG 1995).

Eredményeink egyértelműen azt mutatják, hogy a kísérletben vizsgált magánhangzóknál az $F_1 : f_0$ hangolási tendencia megjelenése minden énekes esetében az F_1 beszédbeli értéktől függött, azaz az F_1 hangolásának alsó határa a magánhangzók nyíltsági fokával vagy a képzésükhöz szükséges vízszintes nyelvmozgással és/vagy az állkapocsnyitás szögével volt

összefüggésben. A legalacsonyabban, az f'' zenei hang környékén, a zárt, első nyíltsági fokú vagy felső nyelvállású magánhangzók F_1 -formánsát hangolták az énekesek az f_0 értékére, melyet a második nyíltsági fokú, középső nyelvállású hangok követtek az f'' zenei hang környékén, majd pedig a harmadik és negyedik nyíltsági fokú, alsó és legalsó nyelvállásfokú hangok a h'' zenei hang magasságában.

A jelen vizsgálat eredményei nagyrészt egyetértésben vannak az eddigi legtöbb magánhangzó-minőséget együttesen vizsgáló tanulmány akusztikai eredményeivel (JOLIVEAU et al. 2004), ám az idézett kísérletben talált egyik jellegzetességet saját eredményeinkben nem tapasztaltuk: a jelen kísérletben ugyanis az énekesek nem érték el az $F_1 : f_0$ stratégia felső határát az ajakkerekítéses magánhangzóknál (ezt JOLIVEAU és munkatársai eredményei alapján 800-1000 Hz körül vártuk). Ennek nyomán megkockáztatható az a feltételezés, hogy a két vizsgálat közti különbség egy, a formánshangolás mint énektechnikai stratégia és az énekes hangtartománya közti összefüggésre utalhat a következő okfejtés szerint.

GARNIER és munkatársai (2010) vizsgálatuk alapján úgy ítélik meg, hogy az $F_1 : f_0$ hangolásának generikus és magánhangzó-független (inkább az ember anatómiai és fiziológiai működései által meghatározott) felső határa van, mégpedig 1100 Hz körül. Ettől az alapfrekvenciától kezdve ugyanis az $F_1 : f_0$ hangolás nem lehetséges (mert nem lehetséges az állkapocsszög további nyitása sem). Ezzel összefüggésben a szerzők a vizsgálatban 1100 Hz fölött más stratégiák, például az $F_2 : f_0$ hangolás megjelenését találták, de természetesen csak abban az esetben, ha az 1100 Hz fölötti frekvenciatartomány még az énekes által használt hangtartományhoz tartozott. JOLIVEAU és munkatársainak (2004) eredményei alapján úgy tűnik, hogy az ő vizsgálatukban részt vett énekesek közül többen is elérték az 1042 Hz alapfrekvenciát (vagy akár magasabb alapfrekvenciákat is) az f_0 emelésekor, tehát feltehető, hogy ezen énekesek hangtartományának felső határa magasabb volt, mint a hangolás szempontjából kritikusnak tűnő 900 Hz. A jelen vizsgálatban azonban úgy tűnt, hogy a 988 Hz alapfrekvenciájú h'' hang elérésével megközelítőleg a hangtartományuk felső határát is elérték az énekesek – amennyiben a hangtartományt definiálhatjuk a magas hangok viszonylag tartós éneklésre való képesség meglétével. Ha tehát feltesszük, hogy a jelen vizsgálat adatközlőinek hangtartományát megközelítőleg lefedték a h'' hangig terjedő alapfrekvenciákkal – szemben JOLIVEAU-ék adatközlőivel, akiknél magasabbra kiterjedő alaphangtartomány tételezhető fel –, két következtetésre juthatunk.

Az egyik következtetés szerint a jelen vizsgálat adatközlőiként szolgáló énekesek a h'' zenei hang, azaz 988 Hz alapfrekvencián elérték a hangtartományuk (fiziológiai paraméterek által befolyásolt) felső határát, így **nem volt szükségük** más stratégiák (például az $F_2 : f_0$ hangolás) ismeretére és alkalmazására. Egyúttal ugyanakkor ez azt is jelentette, hogy képesek voltak kiterjeszteni az $F_1 : f_0$ hangolási stratégiát a hangtartományuk legmagasabb alapfrekvenciájáig még az ajakkerekítéses hangzók esetében is. A másik lehetséges értelmezés szerint viszont az is feltételezhető, hogy maguk a formánsstratégiák befolyásolják a hangtartomány kiterjedtségét, tehát az énekesek éppen azért érték el a hangtartományuk felső határát, mert **nem tudnak** más, az $F_1 : f_0$ -t helyettesítő stratégiákat alkalmazni (például az $F_2 : f_0$

hangolást). Azt, hogy a két felvetés közül melyik helytálló, tehát hogy az énekes által birtokolt formánsstratégiák következnek-e a más faktorok által meghatározott hangtartományból (és annak kiterjedéséből), vagy pedig inkább a formánsstratégiák determinálják-e a hangtartományt, e helyütt – célzott vizsgálat híján – nem áll módunkban eldönteni, de további vizsgálatokra feltétlenül érdemesnek tartjuk a kérdést.

Az énekelt magánhangzók F_2 -értéke a magánhangzók képzéséhez szükséges vízszintes nyelvműködés függvényében változott: az eredmények alapján úgy tűnik, hogy e tekintetben a veláris magánhangzók F_2 -je a palatális hangoknál érzékenyebb az F_1 hangolására (és az azt előidéző artikulációs változásokra). Az eredmények szerint a hátul képzett magánhangzók F_2 -értéke az $F_1 : f_0$ hangolás megjelenésétől kezdve jelentősen növekedni kezdett az alaphéremencia emelésével, míg az elől képzett hangzóknál csak kisebb mértékű változásokat találtunk (az elől képzettek esetében az F_2 értéke valamelyest csökkent a legmagasabb alaphéremenciákon). Mindezzel kapcsolatban – JOLIVEAU és munkatársai (2004) magyarázatát is kölcsönözve – azt feltételezzük, hogy az F_2 -t érintő változás nem szándékos, az F_2 közvetlen hangolását célzó stratégia eredménye, hanem az F_1 hangolását kísérő törvényszerű következmény, ahogy ugyanis az énekes növeli az állkapocs nyitásszögét az f_0 emelése érdekében, a nyelvhat által okozott szűkület is törvényszerűen egyre nagyobb lesz (a veláris vagy a palatális régióban). Mivel a nyelv és az állkapocs bizonyos mértékig külön-külön is működtethető artikulációs szervek, az állkapocsnyitás növelése miatt egyre kisebb (a nyelvhat és a szájpád között található) szűkület a nyelv emelésével egy darabig kompenzálható, hiszen a nyelvhat magasságának intenzív növelésével az állkapocs vízszintes helyzetének csökkentése mellett is fenntartható a szűkület közel eredeti mérete. Ugyanakkor LINDBLOM és SUNDBERG (1971) egy artikulációs modellen alapuló vizsgálatának eredményei szerint ez a kompenzáció csak egy bizonyos mértékű állkapocsnyitásszög eléréséig lehetséges. Ha ugyanis a metszőfogak távolsága eléri a 23 millimétert, a nyelv már nem képes megfelelően kicsi szűkületet létrehozni – legalábbis nem a veláris területen –, így az eredetileg hátul képzett magánhangzó centralizálódni kezd, azaz előrébb tolódik az akusztikai magánhangzóterben, mert második formánsának értéke növekszik. Az állkapocs nyitásszögének növelése tehát artikulációs értelemben egy adott határértéken túl törvényszerűen eredményezi az ejtési szándék szerint hátul képzett hangzók centralizálódását is. Bár LINDBLOM és SUNDBERG (1971) szerint az állkapocsnyitásszög növelése a palatális hangzók ejtésére is hatással van, azok a velárisokhoz képest csak jóval nagyobb állkapocsnyitás esetén kezdenek centralizálódni. Ennek eredményeképpen, ha az állkapocsnyitás szöge nő, a velárisok F_2 -értéke már viszonylag alacsony állkapocsnyitásfoknál növekedni kezd, ám a palatálisok F_2 -értéke csak jóval nagyobb nyitásszögnél változik. LINDBLOM és SUNDBERG (1971) akusztikai modellje szerint tehát az állkapocsnyitásszög növelése a magánhangzók ejtésében az elől és hátul képzett magánhangzók akusztikai konvergálását eredményezi, ami magyarázza a kísérletben kapott akusztikai adatokat is.

Az alaphéremencia emelését kísérő, az F_1 -et és F_2 -t érintő változások „végeredményeként” a legmagasabb alaphéremencián az F_1 az első harmonikussal, tehát az f_0 -lal, az F_2 pedig

a második harmonikkussal, tehát a $2f_0$ -lal esett egybe (a formánsok tehát ezt a két harmonikust erősítették). Ez viszont nem jelent mást, mint a magánhangzók eltolódását az akusztikai magánhangzótérben a legnyíltabb centrális pozíció, azaz az /a:/ magánhangzó pozíciója felé.

A manuális inverz szűrés korábban már eredményesnek bizonyult a formánsmeghatározásban 500 Hz alatti férfi éneklés esetében (vö. SUNDBERG et al. 2013) és magas alapprofrekvenciájú női éneklés esetében (legalábbis az /a/ magánhangzónál, vö. HERTEGÁRD–GAUFFIN 1993). A jelen vizsgálatban az inverz szűrést SUNDBERG és munkatársainak (2013) vizsgálatához hasonlóan HENRICH és munkatársainak (2004) eredményeire is alapozva egy elektroglottográf által rögzített elektroglottogram segítségével (még pontosabban az elektroglottogram deriváltjának segítségével) végeztük el. Az inverz szűrés során kapott eredmények megbízhatóságát egy külső gerjesztés (elektromos gége) segítségével rögzített hanganyag elemzése révén is biztosítottuk, hiszen az elektromos gége lehetővé tette az énekesek toldalékcső-rezonanciáinak a zöngétől és a zöngé alapprofrekvenciájától független detektálását. A vizsgálati eszközök ezen új kombinációja hatékony módszernek tűnik a formánsmeghatározásban még magas alapprofrekvenciákon is. A jelen vizsgálat további újdonsága, hogy először kísérelte meg rendszerszinten elemezni a magas alapprofrekvenciájú éneklést kísérő akusztikai változásokat egy nyelv összes magánhangzójának szisztematikus vizsgálata révén.

A következő fejezetben arra teszünk kísérletet, hogy az akusztikai adatok birtokában a hallási észlelés nem lineáris sajátosságait is figyelembe véve elemezzük a magas alapprofrekvenciákon énekelt magánhangzókat jellemző észlelési tendenciákat. Ismételten fel kell hívunk a figyelmet, hogy arra a kérdésre, mely azt firtatja, hogy milyen tendenciák jellemzik a nem az ejtési szándék szerinti azonosításokat a magas alapprofrekvenciájú szopránéneklésben és miért, korábban még nem született kielégítő válasz.

7. A szopránénekülésben megjelenő magánhangzók észlelése az akusztikai szerkezet tükrében

A korábbi szakirodalomban meghatározónak látszó feltételezés szerint az alapfrekvencia emelésével az éneklésben az $F_1 : f_0$ hangolás következtében egyre nyíltabban ejtett magánhangzókat olyan beszédhangokkal tévesztik össze hallgatók, melyek a beszédben nyíltabb ejtésűek, illetve magasabb F_1 formánssal képzettek, mint az ejteni szándékozott beszédhang (lásd pl. SCOTTO DI CARLO–GERMAIN 1985; HOLLIEN et al. 2000). A kötet eddigi fejezeteiben ugyanakkor a szakirodalom részletes elemzésével, valamint saját vizsgálatainkkal azt igyekeztünk bemutatni, hogy e feltételezések mögött valójában nem áll megbízható és erős empirikus bizonyíték. Mi több, egyes saját eredményeink (lásd a 2. és 3. fejezeteket) e feltételezésekkel egyenesen ellentétesnek mutatkoztak, hiszen több vizsgálatunkban is a zárt ejtésű magánhangzók jó arányú azonosítását találtuk közepesen magas alapfrekvenciákon, valamint azt tapasztaltuk, hogy a magas alapfrekvenciákon ejtett hangzókat nagy számban inkább zárt hangzókként azonosítják a hallgatók, nem pedig nyíltakként.

Mivel korábban nem álltak rendelkezésünkre megbízható rezonanciaadatok az egyes hangzók ejtésekor jellemző formánsokról, nem vizsgálhattuk adatainkat részletesebben például azon elmélet kereteiben, mely a zártságészleletet pszichoakusztikai kísérletekre alapozva az F_1 és f_0 tonotópiás távolságából (az f_0 és $2f_0$ távolságával, tehát az alapfrekvenciával összefüggésben) vezeti le. Az előző fejezetben bemutatott kísérletben azonban az énekelt magánhangzóról formánsadatokat nyertünk három szoprán énekes ejtésében, így lehetőségünk nyílik a percepciósi adatokat az (érzetileg transzformált) rezonanciák fényében értelmezni. Ez tehát a jelen fejezetben bemutatásra kerülő elemzés célja. Ehhez egy, az előzőekhez nagyban hasonló észlelési tesztet végzünk.

Az észlelési adatok elemzéséhez az itt következőkben felhasználjuk TRAUNMÜLLER (1981) modelljét, mely a magánhangzók észlelt zártságát elsősorban az F_1 és az f_0 (tonotópiás) távolságából eredezteti. Az elméletről átfogóan már az 1.1.2.2. *alfejezet*ben írtunk, a hozzá kapcsolódó kísérletet és az elméletnek a jelen kérdéskörre vonatkozathatóságát pedig az itt következő 7.1. *alfejezet*ben ismertetjük. A kísérlet előfeltevései a következők.

A kötet korábbi fejezeteiben bemutatott vizsgálatok tanulságaira alapozva azt feltételezzük, hogy 900 Hz fölött a legtöbb énekelt magánhangzó-megvalósulásra /a:/ válasz érkezik (az ejteni kívánt hangzó minőségétől függetlenül), alacsonyabb alapfrekvenciákon azonban a magánhangzók az alapfrekvencia emelésével egyre bizonytalanabb minősége miatt jóval változatosabb lesz a válaszok megoszlása. A hallási észlelés tonotópiás szerveződésén és az F_1 és f_0 távolságának tonotópiás viszonyán alapuló elméletből kiindulva azt várjuk, hogy

a modell szerint kritikus 370 Hz alatt (azaz a kísérletben vizsgált 349 Hz alapprofrekvencián, a f' zenei hangon és az alatt), az $F_1 : f_0$ hangolás megjelenése előtt a hangzók zártságának észleletét az F_1 és f_0 észleleti távolsága határozza meg. Ugyanakkor azt is feltételezzük, hogy azon hangzók esetében, melyek ejtések az énekesek alkalmazzák az F_1 -nek az f_0 -ra hangolását, TRAUNMÜLLER tonotópiás elméletének értelmében a hallgatók nagy arányban azonosíthatnak zárt hangzókat is, illetve megjelenhet az egyes ejteni szándékozott középső nyelvéllású (második nyitásfokú) beszédhangok zártakként való felismerése is.

A vizsgálatban kapott percepciók eredmények itt bemutatásra kerülő elemzési kísérlete az énekelt beszédhangok kutatásában újdonságnak számít, miközben segítségével módunk nyílik magyarázatot találni a jelen kötet korábbi fejezeteiben ismertetett vizsgálatok egyes, akkor magyarázat nélkül maradt észlelési tendenciáira is: a zárt ejtésű hangzók nagyarányú azonosítására magas alapprofrekvenciákon, valamint a tévesztési tendenciákban az egyes hangzók zárt ejtésű hangzókként való azonosítására.

7.1. Traunmüller modellje és a modell lehetséges predikciói az énekelt magánhangzók azonosítására

TRAUNMÜLLER (1981) vizsgálatait olyan osztrák dialektusok beszélőivel készítette, mely dialektusokban ötféle fonológiailag kontrasztív nyíltsági fok különül el. Az első fok a legzártabb, míg az ötödik fok a legnyitottabb hangzókat jelöli. Ezek rendre a következők: 1. /i u y/; 2. /e o ø/; 3. /ɛ ɔ œ/; 4. /æ ɒ œ/; 5. /a/. TRAUNMÜLLER ezen hangzók segítségével vizsgálta azt a kérdést, hogy milyen szerepet játszik az f_0 és az F_1 tonotópiás, azaz érzeti távolsága a magánhangzók zártságának észleletében.

A modell a már CHISTOVICH és LUBLINSKAJA (1979) által kimutatott 3 Bark terjedelmű másodlagos kritikus sávon alapszik (elsődleges kritikus sávnak az 1 Barkot tekintjük, mely azt a frekvenciafüggő frekvenciasávot jelöli, amely – az alaphártya felépítéséből következően – a fül frekvenciafeldolgozásának legkisebb egysége). A másodlagos, 3 Bark terjedelmű kritikus sáv integrációs határként értelmezhető: azon két akusztikai prominencia, mely 3 Bark észleleti távolságnál kisebb távolságra helyezkedik el egymástól, az észleletben egyetlen, frekvenciájában a két prominencia frekvenciaértéke közötti (azok amplitúdójával súlyozott) „gravitációs középpontba” (súlypontba), azaz egyetlen észleleti prominenciába integrálódik.

TRAUNMÜLLER (1981) modellje szerint abban az esetben, ha két harmonikus (tehát például az f_0 és $2f_0$) között, melyek egy rezonancia (pl. F_1) közelébe esnek, kisebb a távolság, mint 3 Bark, a két harmonikus mint akusztikai prominencia integrálódik, mégpedig a közeli (a példában az F_1) rezonanciának az értékéhez közeli prominenciában. Mivel 370 Hz alatt az f_0 és a $2f_0$ távolsága mindig kisebb, mint 3 Bark, ezért a 370 Hz alatti alapprofrekvenciákon az f_0 és a $2f_0$ minden esetben (magánhangzó-minőségtől függetlenül) egyetlen érzeti csúcsot

eredményez. Mivel pedig az f_0 és az F_1 az első, második és harmadik nyíltsági fokú (azaz nagyjából a magyarban felső, középső és alsó nyelvvállásúnak nevezett) magánhangzók esetében szintén 3 Barknál közelebb van egymáshoz, így ezen hangzók esetében az f_0 és F_1 tartományán mindig csak egyetlen érzeti prominencia jelenik meg, amely tehát (megközelítőleg) az F_1 értékét veszi fel.

370 Hz alatt a TRAUNMÜLLER kísérletében megkülönböztetett első három nyíltsági fokú magánhangzó úgy különül el egymástól, hogy az első és második, valamint a második és harmadik nyíltsági fokú magánhangzók közti különbséget az F_1 és f_0 távolsága – azaz a fentiek értelmében valójában az F_1 értéke – adja a következőképpen. Ha az F_1 és f_0 különbsége kisebb vagy egyenlő, mint 1,2–1,5 Bark, a hangot első nyíltsági fokúnak észleljük (/u i y/). Ha az F_1 és f_0 különbsége nagyobb, mint 1,2–1,5 Bark, de nem nagyobb, mint 2,5 Bark, a hangzót a második nyíltsági fokba tartozónak észleljük (/e o ø/). Ha pedig az F_1 és f_0 különbsége nagyobb, mint 2,5 Bark, de kisebb, mint 3–3,5 Bark, a hangzót a harmadik nyíltsági fokba tartozónak észleljük (/ɛ ɔ œ/). Mivel az F_1 és f_0 e hangok esetében 370 Hz alatt az F_1 értékének megfelelő frekvenciájú érzeti prominenciába integrálódnak, az iméntieket másként tehát úgy is megfogalmazhatnánk, hogy a zártság észleletét az első, második és harmadik nyíltsági fokú beszédhangok esetében az F_1 értéke határozza meg a beszédben (hiszen a beszéd alaphangfrekvenciája felnőttek esetében általában alacsonyabb, mint 370 Hz).

A negyedik és ötödik nyíltsági fokú (a magyar szakirodalomban legalsó nyelvvállásúnak nevezett) magánhangzók esetében azonban már más a helyzet, hiszen ezekben az f_0 és az F_1 értéke közti különbség nagyobb, mint 3 Bark. A negyedik és ötödik nyíltsági fokú hangokban éppen ezért 370 Hz alaphangfrekvencia alatt az f_0 és F_1 érzetileg nem integrálódnak, azaz ezekben a magánhangzókban az F_1 és f_0 frekvenciatartományán nem egy, hanem két érzeti prominencia jelenik meg. Ezzel összefüggésben az elmélet szerint a negyedik nyíltsági fokú hangzókat az F_1 és f_0 között lévő 3 Barknál nagyobb távolság definiálja az észlelés számára: ha az F_1 és f_0 közti különbség nagyobb, mint 3 Bark, de kisebb, mint 5,5–6 Bark, a hangzót negyedik nyíltsági fokú magánhangzónak észleljük (/æ ɒ œ/), míg ha az F_1 és f_0 különbsége ennél is nagyobb, az észleletben a legnyíltabb /a/ jelenik meg. A 370 Hz alaphangfrekvencia alatt ejtett (negyedik és ötödik nyíltsági fokú) magánhangzók további jellemzője az is, hogy ezekben az első (az f_0 -hoz eső) érzeti csúcs kevésbé prominens, mint a második ($2f_0$ -hoz eső) csúcs, ugyanis a magas frekvenciájú F_1 ez utóbbit erősíti.

Míndez azonban hangsúlyozottan csak a 370 Hz alaphangfrekvencia alatt ejtett hangzókra igaz. 370 Hz alaphangfrekvencia fölött ugyanis már az f_0 és az azt követő első felharmonikus ($2f_0$) közti távolság is átlépi a kritikus integrációs (3–3,5 Barkos) határt. A kritikus határ átlépése a fentebb okfejtés értelmében azt eredményezi, hogy a 370 Hz alaphangfrekvencia fölött ejtett beszédhangok észleletében minden esetben két érzeti csúcs jelenik meg a kérdéses tartományon (az f_0 és a $2f_0$ helyén), függetlenül az F_1 értékétől, tehát még akkor is, ha az első három nyíltsági fokba tartozó hangzóról van szó. Könnyen belátható, hogy a modell szerint ez egyúttal azt is jelenti, hogy 370 Hz alaphangfrekvencia fölött a beszédhangok a 370 Hz alaphangfrekvencia alatt ejtett legnyíltabb (/æ ɒ œ/, illetve /a/) hangzókra kezdenek hasonlítani, hiszen a két ér-

zeti prominencia helye a 370 Hz alapprofrekvenciánál magasabb alaphangon ejtett hangokban a 370 Hz alapprofrekvencia alatt ejtett legnyíltabb beszédhangokéval (kvázi) megegyező.

A modell tehát azt jósolja, hogy a különböző magánhangzók megvalósulásai 370 Hz alapprofrekvencia fölött – a hangzók észleleti prominenciáitól vagy valós F_1 -értékétől függetlenül – majdnem minden hangzóra ugyanazt az „észleleti eredményt” adják. A „majdnem” azonban nagyon fontos, hiszen ha ténylegesen és pontosan a modell predikcióit követve próbáljuk értelmezni a 370 Hz felett ejtett beszédhangok azonosítását, akkor azt találjuk, hogy a 370 Hz alapprofrekvencia alatt ejtett legnyíltabb beszédhangokhoz valójában csak a 370 Hz alapprofrekvencia fölött ejtett második és harmadik nyíltsági fokú hangzók válnak egyértelműen hasonlósá. Mivel a második és harmadik nyíltsági fokú beszédhangokban az F_1 frekvenciája magasabb, 370 Hz alapprofrekvencia fölött inkább a $2f_0$ -hoz közelebb, ezért ezekben a két érzeti prominencia (f_0 és $2f_0$) közül a második csúcs lesz intenzívebb, prominensebb – ez pedig valóban a 370 Hz alapprofrekvencia alatt ejtett legnyíltabb hangzókhoz hasonló mintázat.

Az első nyíltsági fokú (/i u y/) hangzók esetében azonban más a helyzet, mint fentebb írtuk, és ezekre a modell valójában nem is fogalmaz meg egyértelmű jóságot. Az első nyíltsági fokú beszédhangokban (370 Hz fölötti alapprofrekvencián ejtve) ugyanis nem a második, hanem az első, az f_0 közelébe eső prominencia lesz intenzívebb, prominensebb, hiszen az F_1 itt az f_0 -hoz esik közel, azt erősíti. Ezt a prominenciamintázatot azonban a legnyíltabb hangzókánál sosem tapasztaljuk „normál”, 370 Hz alapprofrekvencia alatti ejtésben. Így tehát az első nyíltsági fokú (zárt, felső nyelvállású) beszédhangok – a többi nyíltsági foktól eltérően – 370 Hz fölött nem alakulnak az észleletben törvényszerűen a legnyíltabb ejtésű hangzókká. Akkor vajon mégis mi történik ezekkel a zárt ejtésű hangzókkal 370 Hz alapprofrekvencia fölött az észlelésben? Erre a kérdésre TRAUNMÜLLER (1981) modellje nem fogalmaz meg egyértelmű predikciókat, az ugyanazon tanulmányban bemutatott kísérleti eredményei alapján azonban az látszik, hogy az első nyíltsági fokú hangzók a teljes általa vizsgált alapprofrekvencia-tartományon (megközelítőleg 700 Hz-ig) felismerhetőek voltak, és elkülönültek a többi hangzótól. Mindebből következően tehát, érdekes módon, a tonotópiás modell szerint a 370 Hz-nél magasabb alapprofrekvenciákon csak első (zárt, felső nyelvállású) és negyedik/ötödik (alsó és legalsó nyelvállású) nyíltsági fokú magánhangzókat, azaz csak a legzártabb és legnyíltabb beszédhangokat képes azonosítani a beszédpercepció – azokat viszont igen magas alapprofrekvenciáig el tudja különíteni egymástól.

Az énekelt magánhangzók azonosításának kérdését is szem előtt tartva a fentieket a következőképpen foglalhatjuk össze. A modell szerint 370 Hz alapprofrekvencia (mely egyébként az $f_{\#}$ zenei hang alapprofrekvenciája) alatt a zártság észlelete az F_1 és f_0 érzeti távolságának, valamint elsősorban az F_1 értékének függvénye. 370 Hz alapprofrekvencia fölött azonban mindenképpen két prominencia jelenik meg az észleletben, ezért az észleletet és a nyitottság érzetét nem a „teoretikus” (az akusztikai kimenetben nem jól meghatározható) F_1 -nek az f_0 -hoz viszonyított távolsága, hanem az f_0 és a $2f_0$ távolsága befolyásolja elsődlegesen. Mégis, az F_1 frekvenciaértéke még 370 Hz alapprofrekvencia fölött is befolyással van a zártságészleletre, hiszen az F_1 függvényében az f_0 -nál és $2f_0$ -nál megjelenő prominenciák kétféle mintázatot

eredményezhetnek, ami – teoretikusan legalábbis – akár két eltérő észleleti hangzómínőséget is adhat. Az egyik mintázat az, melyben az F_1 a $2f_0$ -hoz (nem pedig az f_0 -hoz) van közel. Ez az, amit a nem első nyíltsági fokú magánhangzóknál tapasztalunk (beszél ejtésben) 370 Hz-nél magasabb alapfrekvencián, és aminek – TRAUNMÜLLER (1981) modelljének predikciója szerint – az eredménye a negyedik nyíltsági foknak megfelelő hangszínű /æ ɒ œ/ (és az egyszerűség kedvéért vegyük ide a TRAUNMÜLLER kísérleteiben ötödik nyíltsági fokúnak minősített /a/ hangot is) megjelenése az észleletben. Percepciósan ugyanis ez a mintázat az /æ ɒ œ a/ 370 Hz alatti megvalósulásaihoz hasonlít a leginkább. A másik mintázat ezzel szemben az, melyben az F_1 nem a $2f_0$ -hoz, hanem az f_0 -hoz van közel, az F_1 itt tehát az f_0 prominenciáját erősíti. Bár ez utóbbi esetre, mint említettük, TRAUNMÜLLERnek (1981) az észlelést modellező teóriája valójában nem ad predikciót, a kísérletes eredmények tanúsága szerint feltehető, hogy az ilyen mintázatot mutató zárt ejtésű (első nyíltsági fokú) hangok megőrzik zárt hangszínüket. Mindehhez adalék, hogy mivel a $2f_0$ és a $3f_0$ (valamint a szekvenciálisan egyre magasabb vagy nagyobb sorszámú felharmonikusok) távolsága sosem nagyobb 3 Barknál, így a $2f_0$ felett megjelenő prominenciákat ismét csak a formánsok frekvenciái, nem pedig a felharmonikusok távolsága befolyásolja még 370 Hz alapfrekvencia fölött is.

A (magas alapfrekvencián) énekelt magánhangzók észlelésének vizsgálatában egyfelől az alapfrekvencia változásának, és így f_0 és $2f_0$, valamint az F_1 és f_0 távolságának hatásaival, másfelől pedig az $F_1 : f_0$ hangolás befolyásával is számolnunk kell, mely akusztikai stratégiát akkor várjuk, ha az f_0 frekvenciaértéke magasabb, mint az F_1 beszédbeli megjelenésének értéke. Mindennek figyelembevételével az énekelt magánhangzók észlelésének alakulásáról a következőt mondhatjuk.

Az alapfrekvencia emelkedése az $F_1 : f_0$ hangolás megjelenése előtt és 370 Hz alapfrekvencia alatt azt eredményezi, hogy az F_1 és f_0 távolsága csökken, ami a tonotópiás modell predikciója szerint az egyes beszédhangok zártabbként azonosításához vezethet akkor, ha az F_1 és f_0 különbsége átlépi a nyíltsági fokokat meghatározó határértékeket. Ha már itt, 370 Hz alapfrekvencia alatt fellép az $F_1 : f_0$ hangolás (ami feltehetőleg csak az első nyíltsági fokú, alacsony F_1 formánssal rendelkező /i: y: u:/ hangzóknál várható), az feltehetőleg nem okoz változást a zártágészleletben, hiszen csak tovább csökkenti az F_1 és f_0 percepciósi távolságát (azaz csökkenti az észlelt F_1 értékét), és zárt észleletet eredményez. A magasabb beszédbeli F_1 formánssal rendelkező, a második, harmadik vagy negyedik nyíltsági fokba tartozó magánhangzók esetében 370 Hz alapfrekvencia alatt pedig nem várunk $F_1 : f_0$ hangolást.

370 Hz alapfrekvencia fölött elméletben két esetet különíthetünk el: azon magánhangzóknak az esetét, amelyeknek az első formánsát az énekesek a megemelt f_0 -ra hangolják (melyek F_1 -értéke a beszédben alacsonyabb), valamint azon magánhangzókét, amelyekét nem. Azokban az esetekben, amikor 370 Hz fölötti alapfrekvencián nincs $F_1 : f_0$ hangolás, az észleletben elkülönülő két (az f_0 -hoz és a $2f_0$ -hoz eső) prominenciából a második az intenzívebb (hiszen ezt erősíti a magas frekvenciájú F_1). Éppen ezért ezeket a hangzókat az azonosítás során az elmélet alapján valószínűsíthető módon a negyedik nyíltsági fokú /a:-nak és /b/-nak felel-

tetik meg a hallgatók. Azon hangok esetében azonban, melyek ejtése során az énekesek 370 Hz alapprofrendencia fölött alkalmazták az $F_1 : f_0$ hangolást, az első, azaz az f_0 -hoz eső prominencia lesz intenzívebb (nem pedig a $2f_0$ -hoz eső). E hangok percepció s státusza a tonotópiás modell alapján nem egyértelműen jósolható. Mivel elméletben ez a prominencia-mintázat nem jellemző a negyedik nyíltsági fokú hangzók 370 Hz alapprofrendencia alatt ejtett megvalósulásaira, viszont hasonlít a 370 Hz alapprofrendencia fölött ejtett első nyíltsági fokú hangzókra jellemző mintázatokra, feltételezhetjük, hogy a 370 Hz fölött ejtett, $F_1 : f_0$ -hangolt beszédhangokat inkább zártakként azonosítják az észlelők, úgy, hogy feltehetően a magánhangzók minősége ezzel együtt meglehetősen bizonytalan. Mindazonáltal ha az f_0 és $2f_0$ prominenciája közti különbség megfelelően kicsi, az első nyíltsági fokú és negyedik nyíltsági fokú válaszok egyaránt valószínűek lehetnek az azonosítás során.

7.2. Anyag, módszer, kísérleti személyek

A vizsgálatban a 6. fejezet akusztikai elemzéseiehez rögzített és az ott akusztikailag elemzett hangmintákon végeztünk percepció s tesztet 21 magyar anyanyelvű, ép halló, egészséges kísérleti személy részvételével (12 nő, 9 férfi, 26 és 45 év közöttiek, életkoruk átlagosan 29 év) egy, a Praat programban (BOERSMA–WEENINK 2009) írt és futtatott szkript segítségével. A három énekes által 6 alapprofrendencián (175 és 988 Hz között, lásd 6. táblázat) felénekel t és beszéd módban ejtett /ɒ a: ε e: i: o: ø: u: y:/ hangzókat a percepció s teszthez némileg manipulálnunk kellett, ehhez a Wavesurfer programot használtuk (SJÖLANDER–BESKOW 2000). A program segítségével kiegyenlítettük az egyes felvételek esetlegesen eltérő hangereősségét, valamint elimináltuk a felvételekből az elektromos zöngé bekapcsolását követő szakaszt (lineáris fokozatos kiűzés effektus segítségével) úgy, hogy az egyes stimulusok megközelítőleg egyenlő (1,8–2 s) hosszúságúak legyenek. A tesztben az ily módon egységített hanganyagot minden adatközlőnek kétszer játszottuk le adatközlőnként eltérő, randomizált sorrendben (így az esetleges fáradás következtében megjelenő hatások eloszlottak a stimulusok között, valamint azt is kontrolláltuk, hogy ne léphessen fel az egymást követő stimulusok szisztematikus egymásra hatása az észleletben).

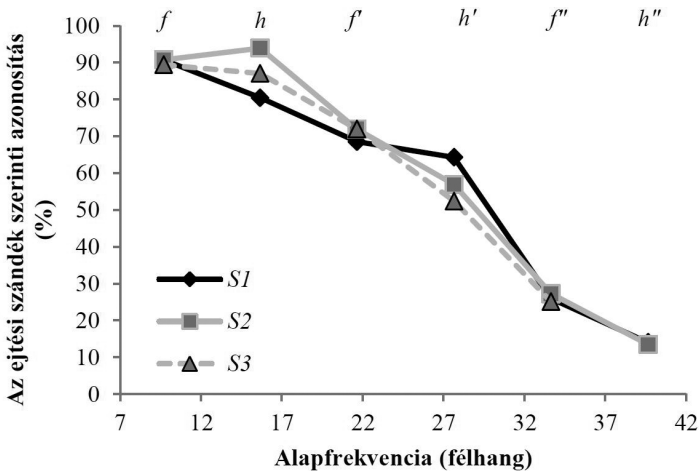
A résztvevőknek a feladat utasítása szerint ki kellett választaniuk a képernyőn ortografikus alakban megjelenő kilenc magyar magánhangzó közül azt, amelyiket hallani vélték. A válaszlehetőségek a hosszú (vagy fonémakategória-váltás nélkül hosszán ejthető) magyar magánhangzók voltak: /ɒ a: ε e: i: o: ø: u: y:/. A teszteket csendes körülmények között vettük fel, a hanganyagokat zárt fejhallgató n keresztül hallották a résztvevők.

Az észlelési teszt adatainak elemzésében figyelembe vettük az akusztikai vizsgálat eredményeit. Bár az akusztikai adatokat érzetileg transzformáltuk, egyúttal azt is hangsúlyozni kívánjuk, hogy a (frekvenciaérzékenység miatt is létrejövő) magas alapprofrendencián ejtett

magánhangzókra jellemző alacsony akusztikai felbontás miatt azt feltételeztük, hogy nem lehet egyszerű összefüggéseket vonni a „nyers” formánsfrekvenciák és az észlelet között.

7.3. Eredmények

Annak megfelelően, hogy a magánhangzók artikulációs és akusztikai célja az alapfrekvencia emelésével változik, azt várjuk, hogy a magánhangzók azonosítása is egyre kevésbé egyezik meg az ejtési szándék szerinti magánhangzó-minőséggel. A három énekes által énekelt kilencféle magánhangzóra kapott, az ejtési szándéknak megfelelő összesített azonosítási arányok e szerint a várakozás szerint alakultak (35. ábra).

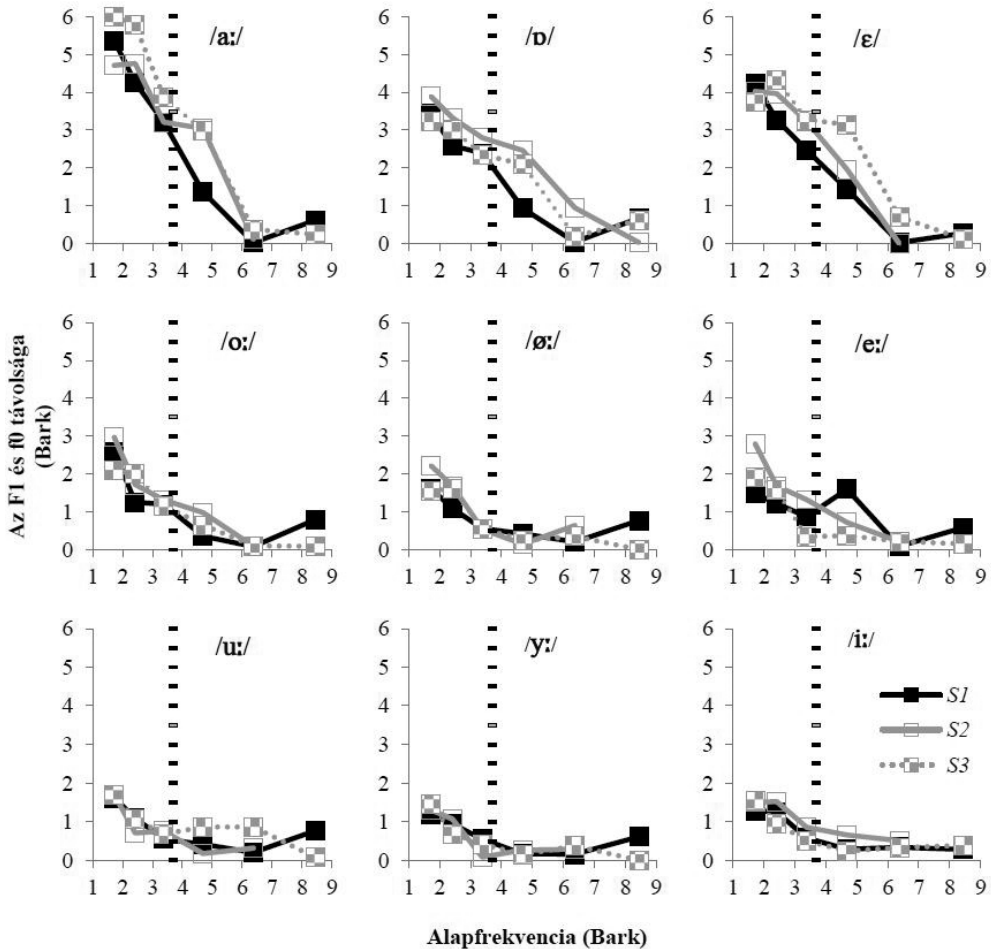


35. ábra: A három énekes (S1, S2, S3) által énekelt magánhangzók azonosítása az alapfrekvencia függvényében

Az ábrán az egyes jelölők egy-egy énekes egy adott alapfrekvencián ejtett összes magánhangzójára kapott (21 adatközlő \times 9 magánhangzó \times 2 ismétlés =) 378 darab válaszból számított százalékos arányt mutatják. Így például azt látjuk, hogy a legalacsonyabb énekelt alapfrekvencián (a zenei f hangon, azaz 175 Hz-en) a mindhárom énekes énekelt magánhangzóra adott válaszok körülbelül 90-90%-a volt az ejtési szándéknak megfelelő, azaz az ejtési szándékkal megegyező hangminőséget megjelölő válasz. Az adatokból kiderül, hogy az énekelt magánhangzókra általánosan mindhárom énekes ejtésében ugyanolyan észlelési tendenciák jellemzőek az alapfrekvencia függvényében: a magánhangzók az ejtési szándéknak megfelelő azonosítása az alapfrekvencia emelésével fokozatosan csökken, míg nem a legmagasabb

h hang (988 Hz alapfrekvencia) elérésével megközelíti a véletlen (11%) szintjét. (Emlékeztetőül: mivel a legmagasabb alapfrekvencián S2 énekes nem tudta végigénekelni az összes kísérleti stimulust, így ezen az alapfrekvencián csak S1 és S3 énekelte magánhangzóit teszteltük).

Az egyes magánhangzókra jellemző azonosítási tendenciák vizsgálatához érzetileg (Bark-skálára) transzformáltuk az előző fejezetben mért és ismertetett F_1 - és f_0 -értékeket. Az F_1 és f_0 Barkban mért különbségét, mely TRAUNMÜLLER (1981) szerint a zártság észleletét befolyásolja (legalábbis bizonyos alapfrekvencia alatt) a 36. ábrán szemléltetjük (*y* tengely) az alapfrekvencia függvényében (*x* tengely).



36. ábra: Az F_1 és f_0 érzeti távolsága (*y* tengely) az alaphangmagasság (*x* tengely) függvényében a három énekes által ejtett kilenc magánhangzóban. A függőleges szaggatott vonalak a TRANUMÜLLER (1981) modellje által jósolt kritikus, 370 Hz alapfrekvenciájú hangmagasságot jelölik

TRAUNMÜLLER (1981) a rezonanciák és az alapfrekvencia tonotopikus viszonyára alapozó modelljének értelmében a különböző nyíltsági fokú magánhangzók elkülönítése szempontjából a 370 Hz alapfrekvencia kritikusnak számít. TRAUNMÜLLER modellje ugyanis elsősorban ez alatt a 370 Hz-es alapfrekvencia feltételez különböző nyíltsági fokú, pontosabban különböző nyíltsági fokúként azonosítható magánhangzókat. Ennek megfelelően, a jelen vizsgálatban az itt vizsgált négyféle nyíltsági fok elkülönítésében a f' (349 Hz) és a h' (494 Hz) zenei alaphangoknak szenteltünk kiemelt figyelmet, valamint elsősorban arra tettünk megfigyeléseket, hogy mi történik ez alatt és e fölött a frekvenciasáv fölött. Emlékeztetőül, a magánhangzók érzeti nyíltságát az F_1 és f_0 távolságából levezető modell szempontjából azért tekinthető fontos határnak a 349 Hz és 494 Hz alapfrekvencia, mert 349 Hz-ről 494 Hz-re lépve az f_0 és $2f_0$ közti távolság 2,99 Barkról 3,75 Barkra nő, tehát átlépi a 3 Barknyi perceptuális integrációs határt, mely alapvetően megváltoztatja az ezeken a frekvenciákon ejtett beszédhangoknak az azonosíthatóságát.

TRAUNMÜLLER elméletéből következően 349 Hz (f') alatt az / υ / és / a / kivételével minden ejteni szándékozott hangzó első formánsának és alapfrekvenciájának frekvenciasávján csak egy prominencia jelenik meg az észleletben – az / υ / és / a / esetében azonban kettő, mert itt az F_1 és f_0 távolsága nagyobb, mint 3 Bark. A legzártabb / i /: u /: y / hangzók esetében tehát ezen az alapfrekvencia-tartományon azt vártuk, hogy ha az F_1 értéke változatlan, és az F_1 és f_0 tonotopikus távolsága nem nagyobb, mint 1,2–1,5 Bark, akkor a hangzókat zártakként, tehát az ejtési szándék szerint ismerik fel a tesztben részt vevők.

A 36. ábra tanúsága szerint bár az / u / esetében a legalacsonyabb énekelt alapfrekvencián 1,2–1,5 Barknál valamivel nagyobb az F_1 – f_0 távolsága, az / u : y : i / első formánsának és az f_0 -nak a távolsága a megadott tartományon belül marad a kritikus, 370 Hz-es alapfrekvenciáig (melyet az ábrán függőleges szaggatott vonal jelez). A 37. ábrán látható azonosítási adatok szerint az észlelésben a modelltől megjósolható tendenciákat tapasztaltuk. Az ábrán ugyanis azt látjuk, hogy míg az / i :- t és / y :- t a kritikus 370 Hz alapfrekvencia alatt, azaz az első három vizsgált alapfrekvencián a legnagyobb arányban első zártági fokú hangzóként (tehát / i :-ként és / y :-ként) azonosították a tesztelők, az / u :- t bár főként az ejteni kívánt hangzó nyíltsági fokának megfelelően észlelték (tehát / u :-ként), mégis kisebb számban, nyitottabb hangzóként (/ o :-ként) is azonosították.

Az akusztikai adatok szerint a kritikus 370 Hz alapfrekvencia alatt a zárt / u : y : i / magánhangzók esetében 349 Hz-en (f') már fellépett az F_1 : f_0 hangolás (33. ábra). TRAUNMÜLLER modellje alapján ugyanakkor azt feltételezhetjük, hogy a hangolás következtében megváltozott akusztikum észlelete ugyanolyan marad, mint ha nem lépne fel hangolás (legalábbis 370 Hz alatt), mert az F_1 : f_0 hangolás következtében az első észleleti prominencia frekvenciája csökken, és ez elviekben csak erősíthet a hangzók első zártági fokú észleletén. A perceptuációs adatok szerint ez a modelltől levezethető predikció helytálló. Az azonosítási adatokban ugyanis azt látjuk, hogy a hangolás nem hoz drasztikus változást az azonosításban, és 349 Hz-en mind a palatális, mind a veláris hangzók esetében gyakorlatilag megismétlődnek az eggyel alacsonyabb alapfrekvencián kapott arányok: a hallgatók az / y : i / magánhangzókat

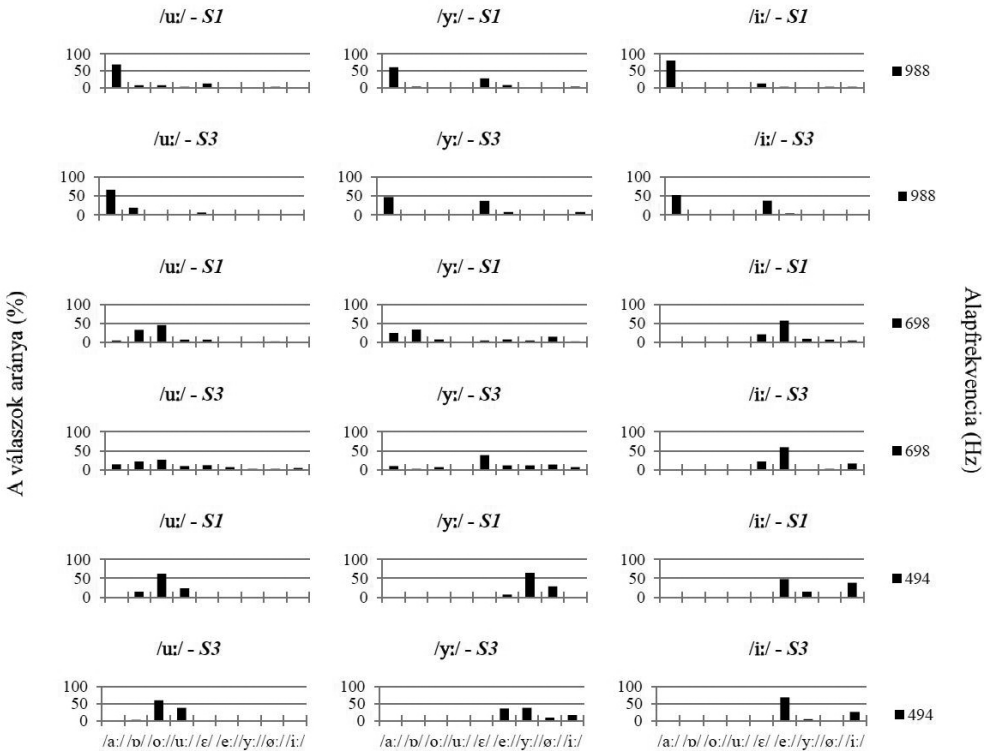
főként első nyíltsági fokú hangokként azonosították, míg az /u/-t nagyobbbrészt /u:/-ként, kisebb számban /o:/-ként.

Akusztikai adataink szerint a nyíltabb /o: ø: e:/ magánhangzók esetében 370 Hz alatt még nem lépett fel az $F_1 : f_0$ hangolás (vö. 33. *ábra*), így ezeknél a hangzóknál csak az alapfrekvencia emelése miatt csökkenő $F_1 - f_0$ távolsággal kell számolnunk mint a zártságészleletet befolyásoló paraméterrel. A modell értelmében e hangok esetében 370 Hz alatt az f_0 és F_1 frekvenciasávján csak egy észleleti csúcs jelenik meg, és e hangok esetében a csökkenő $F_1 - f_0$ távolság zártabb hangzókat eredményez az észleletben. A három második nyíltsági fokú hangzóból csak az /o:/ és /e:/ esetében lehetséges a magyar köznyelvi magánhangzórendszerben olyan módon nyíltabbként észlelni az adott hangzókat, hogy az csak egy jegyben, mégpedig a nyíltsági fokban eltérő hangzókként való azonosítást jelentsen. (Az /ø:/ esetében ugyanis a nyíltabbként azonosításhoz a nyíltabb és ajakréses, tehát két jegyben eltérő /ε/ magánhangzót kellene azonosítani, vagy pedig a két fokkal nyíltabb és hátul képzett /ɒ/-t). Így a nyíltság észlelésének szempontjából a második nyíltsági fokú hangzók közül csak az /o:/ és /e:/ magánhangzókat vizsgáltuk meg mindhárom énekes ejtésében (38. *ábra*).

A három énekes ejtette /o:/ és /e:/ magánhangzók azonosítására az első három alapfrekvencián (370 Hz alatt) nagyon hasonló tendenciák voltak jellemzőek: bár az ejteni szándékozott beszédhangok nagy arányban (még 349 Hz-en is 60-70%-ban) megtartották az ejtési szándéknak megfelelő percepció hangszínüket, az adatközlők az alapfrekvencia emelésével mind az /o/-t, mind az /e/-t egyre nagyobb számban azonosították első zártsági fokú (rendre /u:/ és /i:/) hangzókként. Ugyanakkor az is megfigyelhető, hogy S3 énekes /o:/ hangzója „kilóg a sorból”, hiszen azt 349 Hz-en a hallgatók nagyobb százaléka azonosította az ejtési szándék szerint, mint a többi, azonos nyíltsági fokú beszédhangot, valamint „téveszteni” is inkább egy nyíltabb hangzóra, az /ɒ/-ra „tévesztették” a zártabb /u:/ helyett.

Mivel a harmadik és negyedik nyíltsági fokú /ε ɒ a:/ esetében tendenciózusan ugyanazt az azonosítási mintázatot láttuk mindhárom énekesnél, ezeket a hangokat csak S1 énekes ejtésében szemléltetjük egy, az énekes összes ejteni szándékozott magánhangzójának azonosítását az összes alapfrekvencián bemutató ábrán (39. *ábra*). Mivel az /ε ɒ a:/ esetében 370 Hz alatt még nem mutatkozott az $F_1 : f_0$ hangolás, feltehető volt, hogy a hangzók zártságészleletét ezen a tartományon (a második nyíltsági fokú hangokhoz hasonlóan) csak az F_1 és f_0 távolsága befolyásolja. A 36. *ábrán* szemléltetett akusztikai adatok szerint az /a:/ esetében ez a távolság nem lépte át a 3 Barkos, az /e/ esetében pedig a 2,5 Barkos kritikus határt. Ugyanakkor az /ɒ/ esetében azt látjuk, hogy az F_1 és f_0 távolsága az itt vizsgált második és harmadik alapfrekvencián (azaz a zenei f' és h' hangokon) megközelítette, sőt át is lépte a kritikusnak megállapított (3 Barkos) értéket. Az észlelési adatokban ezzel együtt azt találtuk, hogy az első három vizsgált alapfrekvencián (370 Hz alatt) mindhárom magánhangzót (/ε ɒ a:/) nagy arányban (legalább 70%-ban) az ejtési szándéknak megfelelően azonosították az adatközlők. Ugyanakkor az is kiderült, hogy az /ɒ/ azonosítása az alapfrekvencia emelésével a másik két hangzóénál gyorsabb ütemben csökkent (mígnem 349 Hz-en, a zenei h' hangon elérte a 71%-ot). Az /ɒ/-ra az alapfrekvencia emelésével egyre nagyobb számban érkeztek /a:/ válaszok, amit úgy értel-

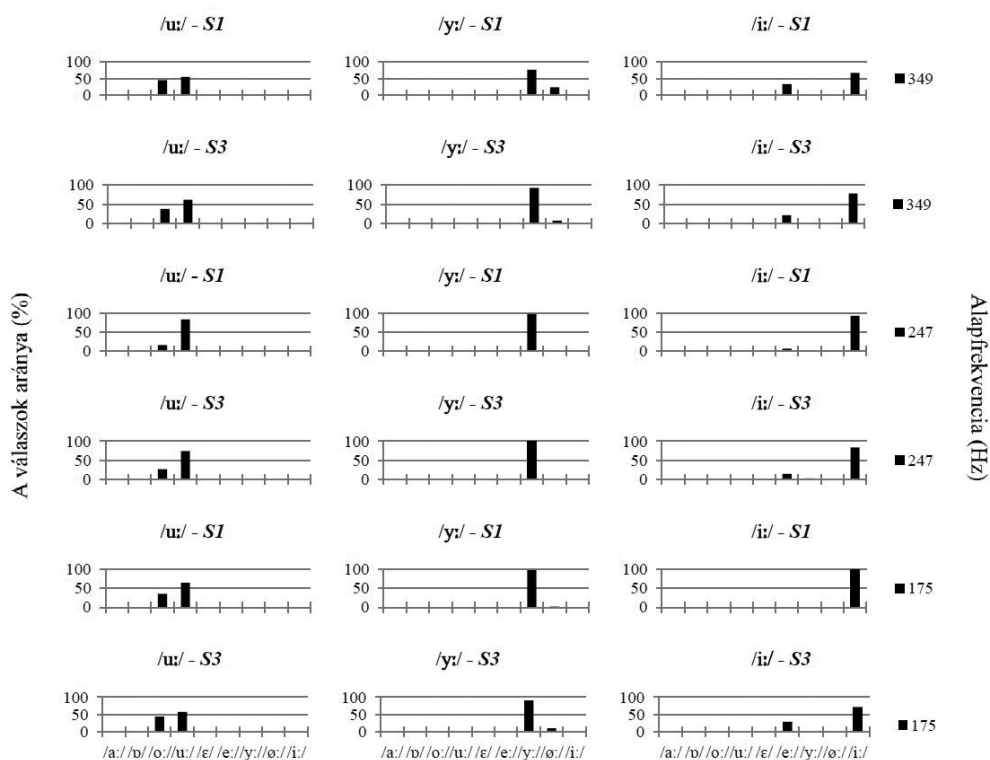
mezzhetünk, hogy az /ɒ/-t ezek az adatközlők valamivel nyitottabbnak érzékelhették (mivel az /a:/ első formánsa magasabb, mint az /ɒ/-é, ezért akusztikailag az /a:/ nyitottabbnak tekinthető, mint az azonos alapprofrecvencián ejtett /ɒ/).



37. a) ábra: Az első nyíltsági fokú, zárt /u: y: i:/ ejteni szándékozott magánhangzókra adott válaszok az f_0 függvényében S1 és S2 énekes ejtésében a 349 Hz-nél magasabb alapprofrecvenciákon

Az integrációs határ átlépése után (ez a jelen kísérletben a zenei f' hang, azaz a 349 Hz alapprofrecvencia elhagyását jelenti az f_0 emelése során) a tonotópiás modell szerint a beszédpercepció minden ejteni szándékozott magánhangzóban két prominenciát azonosít, mégpedig az első két harmonikus, az f_0 és a $2f_0$ frekvenciájának magasságában. Az akusztikai adatokban azt láttuk, hogy az első nyíltsági fokú /i: u: y:/ esetében a 349 Hz fölötti alapprofrecvenciákon már minden alapprofrecvencia és minden énekes esetében fellép az $F_1 : f_0$ hangolás, tehát az f_0 -nál lévő észleleti prominencia felerősödik. Ha a rezonanciák értéke a beszédben mérhetőkhöz képest állandó maradna, TRAUNMÜLLER (1981) modellje és kísérletes adatai alapján azt várnánk, hogy az /i: u: y:/ zártságának észlelete jó arányban (50% fölött) az ejtési szándéknak megfelelő lenne. Az F_1 -nek az f_0 -ra hangolása azonban az f_0 -nál lévő észleleti prominencia

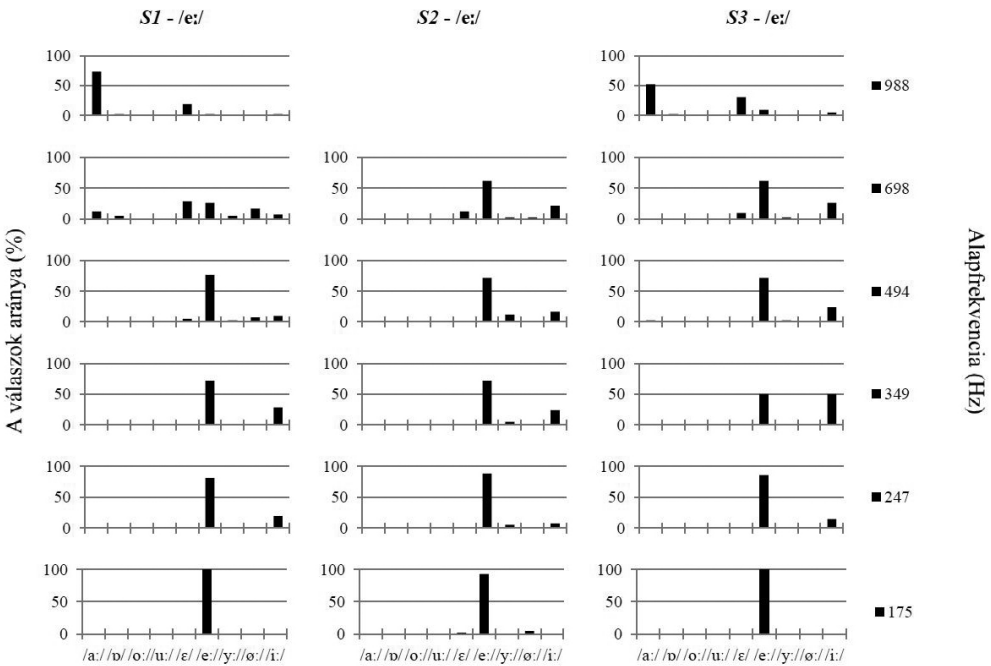
erősödésével, tehát az első nyíltsági fokú hangokra jellemző észleleti mintázatnak a bizonyos mértékű megváltozásával jár, a tonotópiás modellből így az /i: u: y:/ azonosítására jellemző tendenciákat nem jósolhatjuk egyértelműen. A 37. a) és a 37. b) ábrán szemléltetett adataink szerint 370 Hz fölött az alapfrekvencia emelésével az ejtési szándék szerint első nyíltsági fokú /i: u: y:/ magánhangzókat a hallgatók egyre nyitottabb hangokként, végül pedig a legmagasabb zenei hangon, a h'' -n (988 Hz alapfrekvencián) legnagyobb arányban /a:/-ként azonosítják: S1 énekes ejtésében rendre 81%-ban, 69%-ban, valamint 61%-ban, míg S3 ejtésében 52%-ban, 67%-ban, valamint 48%-ban. Mindez pedig egyúttal azt is jelenti, hogy már csak az első nyíltsági fokú hangzókra adott válaszokként is nagy számban jelennek meg az észleletben középső nyelvállású hangok is 370 Hz fölött, ami TRAUNMÜLLER (1981) modelljének a beszéd módra vonatkozó jóslataival nem összeférhető.



37. b) ábra: Az első nyíltsági fokú, zárt u: y: i:/ ejteni szándékozott magánhangzókra adott válaszok az f_0 függvényében S1 és S2 énekes ejtésében a 494 Hz-nél alacsonyabb alapfrekvenciákon

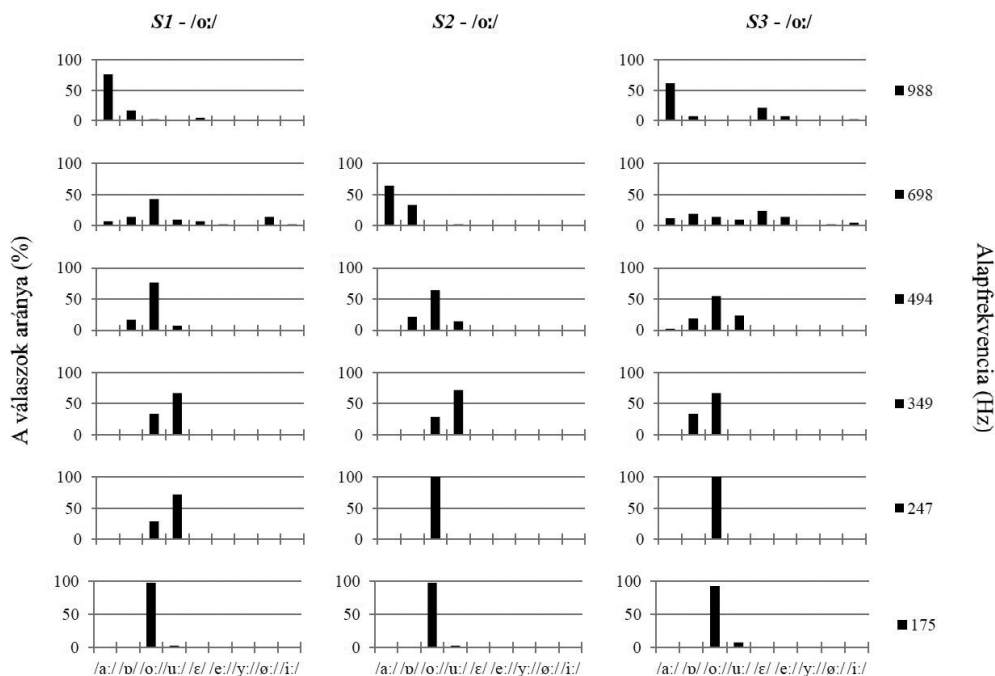
A második nyíltsági fokú /e:/ az f'' -ig (698 Hz-ig) magas, 50% fölötti arányban őrizte meg az ejtési szándék szerinti percepció hangszínét a zártság tekintetében. A 38. ábrán jól látható módon az egyre növekvő számú, az ejtési szándéktól eltérő hangminőségű válaszok között

a leggyakoribb az első nyíltsági fokú /i:/ volt, míg a legmagasabb, *h''* zenei hang (988 Hz alapfrekvencia) elérésével a válaszok legnagyobb részében (az S1 és S3 énekesek ejtésében rendre 74%-ban és 52%-ban) az /a:/-t jelölték meg a hallgatók. Míg az /o:/-ra a kritikus 370 Hz alapfrekvencia alatti legmagasabb alapfrekvenciáig, az *f'*-ig (349 Hz-ig), nagy számban kaptunk /u:/ válaszokat (legalábbis S1 és S2 énekes ejtésében), addig a kritikus alapfrekvencia fölött mindhárom énekes ejtésében ismét megnövekedett (és 50% fölé emelkedett) az erre a hangra adott, az ejtési szándéknak megfelelő válaszok száma. A zenei *f''* hangon (698 Hz alapfrekvencián) az észlelési tendenciák nem voltak egyöntetűek a három énekes ejtette /o:/-kra nézvést. Míg a kísérleti személyek S1 esetében 48%-ban az ejtési szándék szerint azonosították azokat, S2-nél ez az arány jóval magasabb, 64%, S3-nál pedig jelentősen alacsonyabb, mindössze 14% volt (ez utóbbi érték alig valamivel magasabb csak, mint a véletlen szintje, mely ebben a tesztben 1/9 valószínűség, azaz 11% volt). Az /o:/-ra a zenei *f''* hangon (698 Hz alapfrekvencián) kapott, a nem az ejtési szándéknak megfelelő válaszok szintén nem voltak egyöntetűek: míg az /o:/-ra S2 énekes ejtésében a legnagyobb számban /a:/ válaszokat kaptunk (62%-ban), a másik két énekesnél a véletlenhez közeli arányban érkeztek válaszok. A legmagasabb alapfrekvencián, ahogy az /e:/-re is, úgy az /o:/-ra is főként /a:/ válaszokat kaptunk (S1 és S3 énekeseknél rendre 76% és 62%-ban).

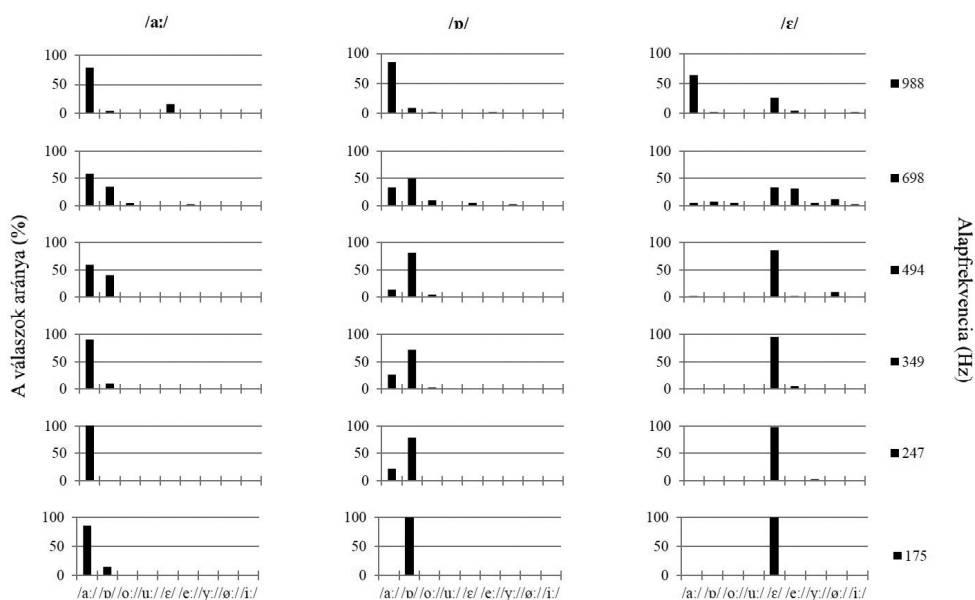


38. a) ábra: A középső nyelvállású, második nyíltsági fokú, a nyitásfok szerint csak egy jegyben eltérő szomszédokkal rendelkező /e:/ ejteni szándékozott magánhangzókra adott válaszok az f_0 függvényében a három énekes ejtésében

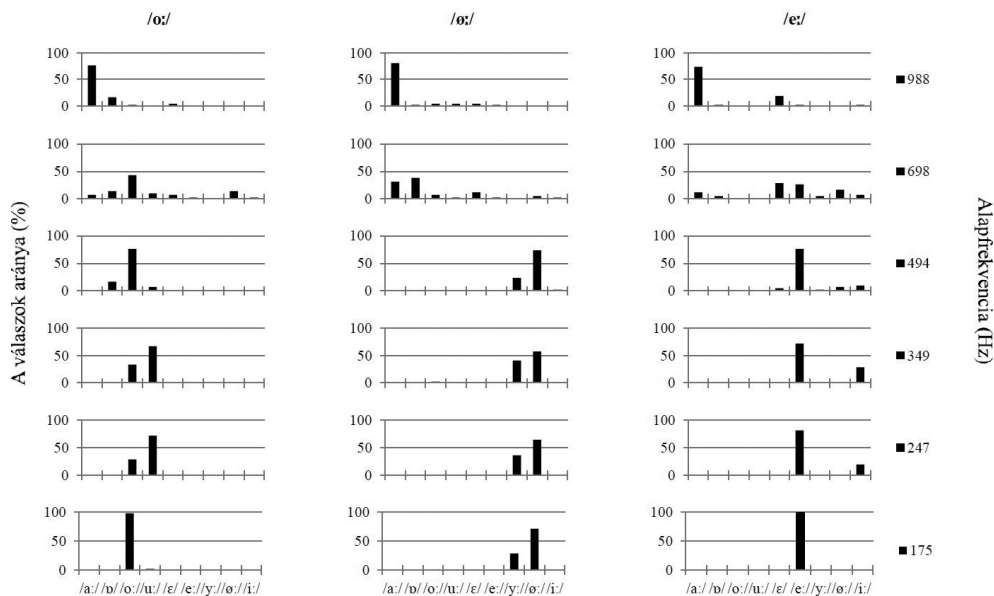
A harmadik és negyedik nyíltsági fokú /ε ɒ a:/ esetében 370 Hz fölött a három énekes ejtette magánhangzókra ismét egymáshoz nagyon hasonló tendenciákat találtunk – így ezeket ismét csak S1 énekes ejtésében szemléltetjük (a 39. a) ábrának az első három sorában található paneleken). Az adatok szerint a harmadik és negyedik nyíltsági fokú magánhangzók a zenei h' hangon (494 Hz alapfrekvencián) még nagymértékben megőrizték érzeti nyíltságukat (az /ε ɒ a:/ rendre 86, 81 és 60%-ban). Ahogyan azonban a többi beszédhangnál is, úgy e hangok esetében is csökkentek az ejtési szándéknak megfelelő azonosítások az f'' alaphangon (698 Hz-en). A legmagasabb, 988 Hz alapfrekvencián azaz a h'' zenei hangon pedig az /ε/, /ɒ/ és /a:/ esetében is a legnagyobb számban /a:/ választ kaptunk (rendre 64, 86 és 79%-ban).



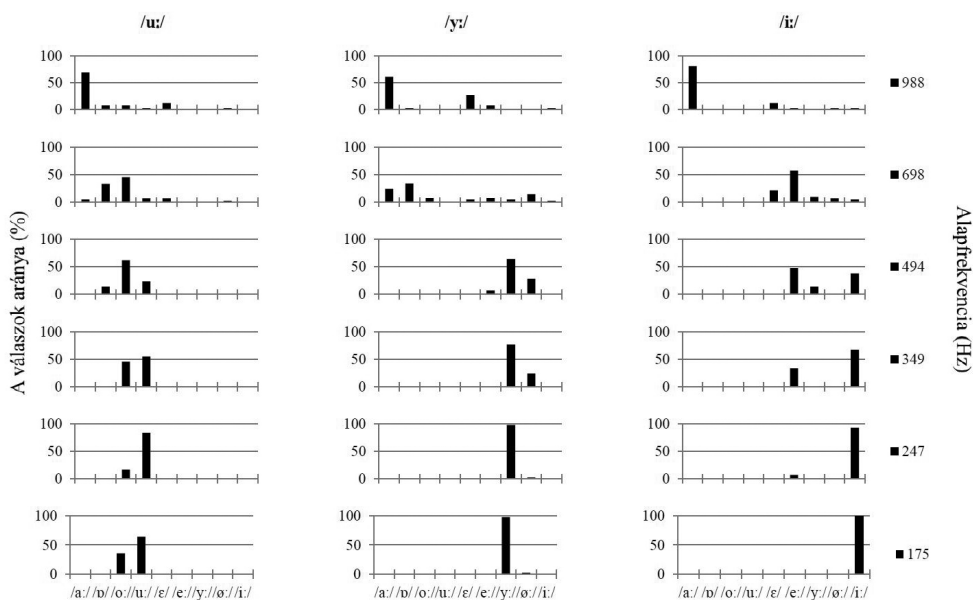
38. b) ábra: A középső nyelvállású, második nyíltsági fokú, a nyitásfok szerint csak egy jegyben eltérő szomszédokkal rendelkező /o:/ ejteni szándékozott magánhangzókra adott válaszok az f₀ függvényében a három énekes ejtésében



39. a) ábra: Az egyes ejteni szándékozott magánhangzókra adott válaszok az alapfrekvencia függvényében S1 énekes énekprodukciónak

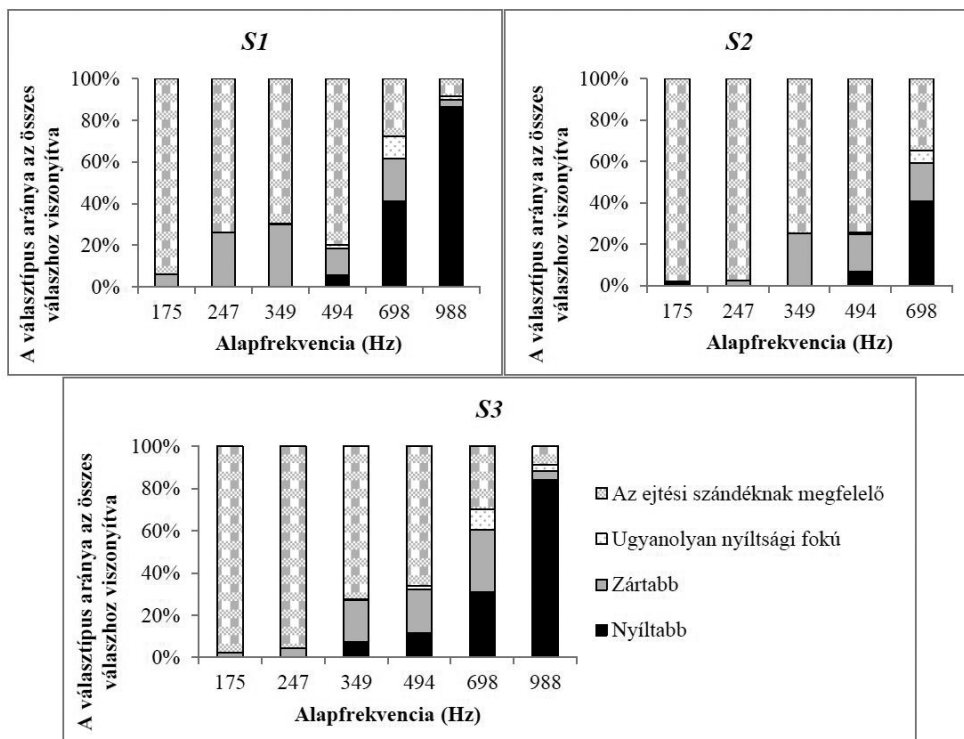


39. b) ábra: Az egyes ejteni szándékozott magánhangzókra adott válaszok az alapfrekvencia függvényében S1 énekes énekprodukciónak



39. c) ábra: Az egyes ejteni szándékozott magánhangzókra adott válaszok az alapfrekvencia függvényében S1 énekes énekprodukciónak

Zárásképp ismét az észlelés valamivel általánosabb tendenciáit elemeztük, mégpedig olyan módon, hogy azok összevethetők legyenek a korábbi, a 2. és 3. fejezetekben ismertett eredményeinkkel is. Kiemeltük a második és harmadik nyíltsági fokú /ε e: o:/ hangzókat (melyek a magyar magánhangzórendszerben nyíltabb és zártabb szomszédokkal is rendelkeznek), és megvizsgáltuk, milyen hangzókként azonosították őket a hallgatók az egyes alapfrekvenciákon az egyes énekesek produkciójában (40. ábra).



40. ábra: A második és harmadik nyíltsági fokú /ε e/; /ø/; /o/ hangzókra az egyes alapfrekvenciákon kapott válaszok megoszlása a magánhangzók nyíltsági fokának függvényében az S1, S2 és S3 énekesnél

Korábbi vizsgálatainkban 500 Hz és 650 Hz közötti (a temperált zenei hangsorok és a nyugati zene hagyományától független, önkényesen választott, lineáris Hertz-értékeit tekintve egymástól ekvidisztáns távolságra lévő) alapfrekvenciákon énekelt magánhangzó-megvalósulásokat elemeztünk, melyekre az ejtési szándéktól eltérő azonosítások esetén dominánsan a hangzók az ejtési szándéknál zártabb minőségű magánhangzókként azonosítása volt jellemző az adott tartományon, az alapfrekvenciától függetlenül. A jelen vizsgálatban ezzel megegyezően azt találtuk, hogy a korábban vizsgált alapfrekvenciák környékén – h' (494 Hz) alaphangon –, illetve az az alatt énekelt magánhangzókra regisztrált észlelési tendenciákat itt is az egyes hangzók zártabb magánhangzókként azonosítása uralja. Ugyanakkor azt is láthatjuk, hogy az alapfrekvencia további emelésével (650 Hz fölött), a kifejezetten magas alapfrekvenciákon a hierarchia élére a hangzóknak az ejtési szándéknál nyíltabb minőségű azonosítása lépett, főként azért, mert – mint azt már említettük – ezeken az alapfrekvenciákon jelentősen megszaporodott az /d/ és /a:/ válaszok száma.

7.4. Következtetések

A fejezetben közölt vizsgálatban szoprán énekesek által magas alapfrekvenciákon énekelt magánhangzók azonosítását vizsgáltuk az ejtés közben jellemző üregi rezonanciák figyelembevételével (mely adatokat a 6. fejezetben bemutatott kísérletben nyertük) nem énekes adatközlők észlelésében. Ehhez, ismereteink szerint elsőként, TRAUNMÜLLER (1981) modelljét alkalmaztuk, mely a magánhangzó észlelt zártságát az F_1 és f_0 tonotópiás távolságából vezeti le.

Az énekelt magánhangzók azonosítására vonatkozó előzetes feltételezéseink részben teljesültek. Míg 988 Hz alapfrekvencián (a zenei h'' hangon) az adatközlők legnagyobb számban /a:/ magánhangzókat azonosítottak az ejtési szándéktól függetlenül, alacsonyabb alapfrekvenciákon jóval változatosabb megoszlást tapasztaltunk a válaszokban, míg általánosan a magánhangzók az ejtési szándék szerinti azonosításában fokozatos csökkenést láttunk az alapfrekvencia emelésével. Bár valóban azt figyelhetjük meg, hogy a hallgatók nagy arányban azonosítanak zárt hangzókat is a magasabb alapfrekvenciákon, az az elvárásunk nem teljesült, hogy az egyes ejteni szándékozott második nyíltsági fokú (középső nyelvállású) beszédhangokat kizárólag vagy legalább dominánsan zártakként ismernék fel a hallgatók. Az általános tendenciák korábbi eredményeinkkel összhangban voltak, beleértve azt az eredményt is, mely szerint a 650 Hz alapfrekvencia alatti hangok esetében a magánhangzóknak a nem az ejtési szándék szerinti azonosításait az ejtési szándéknál zártabb minőségű hangzók azonosítása jellemezte. Korábbi eredményeinket kiegészíti az az itt kapott tapasztalat, hogy az ennél a 650 Hz alapfrekvenciánál magasabb alapfrekvenciákon énekelt magánhangzókra megnövekedett az ejtési szándéknál nyíltabb minőségű hangzóválaszok száma, tehát itt már a SCOTTO DI CARLO és GERMAIN (1985), valamint HOLLIEN és munkatársainak (2000) vizsgálataiban sugalltának inkább megfelelő eredményeket kaptunk.

Az éneklésben a beszédhangok azonosításának tekintetében (az F_1 -nek a 6. fejezet akusztikai elemzésében is megfigyelt változása miatt) a legfontosabb kérdésnek az ejteni szándékozott hangzók nyíltságának azonosíthatósága tűnik, így figyelmünket elsősorban a nyíltság észleletére összpontosítottuk. Ugyanakkor tudvalévó, hogy a hallási észlelési folyamatok frekvenciaérzékenyek, tehát nem lehet egyszerű összefüggéseket vonni a formánsértékek és az észlelet között. Éppen ezért észlelési adatainkat a zártságot az F_1 és f_0 tonotópiás távolságából levezető elmélet keretein belül (TRAUNMÜLLER 1981) az érzetileg transzformált akusztikai adatainkat felhasználva értelmeztük részleteiben.

Az elméletből kiindulva azt vártuk, hogy a modell szerint kritikus 370 Hz alatt (azaz a kísérletben vizsgált 349 Hz (f') alapfrekvencián és az alatt) az $F_1 : f_0$ hangolás megjelenése előtt a hangzók zártságának észleletét az F_1 és f_0 észleleti távolsága határozza meg. Az eredmények megfeleltek a modelltől tehető predikcióknak. A kérdéses alapfrekvenciákon a tesztelők nagy számban azonosították az ejtési szándék szerint az egyes nyíltsági fokú beszédhangokat akkor, ha az F_1 és f_0 távolsága a TRAUNMÜLLER (1981) kísérleteiben az egyes nyíltsági

fokokra jellemzőnek megállapított értékeken belül maradt, ha pedig a távolság a kritikus értékek alá csökkent, a válaszok között – tehát az észleletben – zártabb hangzók is megjelen-
tek. Az $F_1 : f_0$ hangolás megjelenése 370 Hz alatt nem látszott befolyásolni az adatokat – fel-
tehetőleg azért, mert a hangolást csak a legzártabb beszédhangok esetében tapasztaltuk,
a hangolás eredményeként pedig inkább a zárt hangzókra jellemző prominenciamintázat-
hoz hasonló mintázatot feltételezhetünk 370 Hz alapprofrendencia alatt (így a hangolás az észlelt
zárttságon nem változtatott).

A zenei h'' -n, azaz 988 Hz alapprofrendencián talált nagyszámú /a:/ válaszokat az akusztikai
adatok ismeretében (bizonyos mértékig) TRAUNMÜLLER (1981) elméletének keretei közt is
magyarázni tudtuk. Az elmélet ugyanis a 370 Hz fölött két (az f_0 és a $2f_0$ közelébe eső) pro-
minenciával rendelkező hangzókról azt jósolja, hogy azok /a/-szerű hangokként ismerhetők
fel akkor, ha a második észleleti prominencia intenzív (vagy intenzívebb, mint az első promi-
nencia). Bár a zenei h'' hangon ejtett magánhangzók mindegyikénél $F_1 : f_0$ hangolást tapaszt-
altunk, tehát a hangok spektrumának legintenzívebb komponense az f_0 volt, mégis, az észle-
lési adatok alapján feltehetjük, hogy ezek a zenei h'' hangon (azaz 988 Hz alapprofrendencián)
létrejövő prominenciamintázatok nagyban hasonlítottak a negyedik nyitásfokú (/a/-szerű)
hangokra. Akusztikai adatainkban azt láttuk, hogy az f_0 fokozatos emelését nem csak az
 F_1 emelése kísérte a hangolás során, hanem az F_2 változása is: az F_2 ugyanis a velárisok ese-
tében egyre növekvő, a palatálisok esetében pedig enyhén, de csökkenő frekvenciaértéket
vett fel. Az F_2 változásának „végeredménye” a h'' hangon (988 Hz alapprofrendencián) az volt,
hogy az F_2 frekvenciaértéke (gyakorlatilag minden ejteni szándékozott beszédhang esetében)
a zöngé $2f_0$ komponensével esett egybe, tehát erősítette azt. Úgy tűnik tehát, hogy az F_1 és F_2
egybeesése az f_0 és $2f_0$ komponensekkel olyan prominenciamintázatot eredményezett, mely
viszonylag egyértelműen meghatározta a tesztben részt vevő hallgatók számára az /a/-szerű
hangminőséget – szemben az alacsonyabb frekvenciákon tapasztaltakkal, ahol a $2f_0$ -hoz eső
prominencia kisebb intenzitású volt (hiszen nem erősítette azt az F_2), amivel együtt az /a:/ és
/b/ válaszok aránya is alacsonyabb.

Az észlelési eredmények 370 Hz alapprofrendencia fölött (mely alapprofrendencia fölött az $F_1 : f_0$
hangolás a második nyitási fokú, középső nyelvállású hangokban is megjelent) és a zenei h''
(988 Hz alapprofrendencia) alatt nem voltak TRAUNMÜLLER (1981) elméletéből és eredményeiből
jó arányban megjósolhatók. TRAUNMÜLLER (1981) elmélete alapvetően nem ad (vagy legalábbis
nem egyértelmű) predikciót a 370 Hz fölötti alapprofrendencián alacsony F_1 -értékkel ejtett
beszédhangok azonosítására, illetve egyáltalán nem tér ki olyan esetekre, melyek (bármilyen
más módon) az $F_1 : f_0$ hangolt beszédhangok esetéhez hasonlíthatók volnának, így az azokkal
kapcsolatos jóslatokhoz útmutatásul szolgálhatnának. TRAUNMÜLLER kísérleti eredményeiből
kikövetkeztetve azonban mégis azt feltételezhetjük, hogy a 370 Hz fölötti alapprofrendencián
ejtett, alacsony F_1 -értékű hangzóknál az alacsony F_1 , tehát például az $F_1 : f_0$ hangolás eredmé-
nyeként létrejövő akusztikai szerkezet is, az első nyitási fokú (zárt vagy felső nyelvállású)
hangzók észleleti megőrzését segít(het)i, valamint esetleg az egyes, a nem az ejtési szándék
szerint azonosított hangzók zártakként való felismerését is előidézheti.

A modell 370 Hz alapfrekvencia fölött a fentiek mellett a középső nyelvállásfokú hangok csak igen alacsony arányú azonosíthatóságát jósolja a beszédben (TRAUNMÜLLER tesztjében 50% alatti válaszarányt mutatott be 13 vetélkedő válaszlehetőség, tehát 1/13, azaz 8% valószínűség esetén). A jelen vizsgálatban a várakozásokkal ellentétben a középső nyelvállású hangok a jósoltnál jóval nagyobb arányban jelentek meg (akár az ejtési szándékkal megegyező, akár attól eltérő) válaszként 370 Hz alapfrekvencia fölött. Ezen túlmenően pedig nem tapasztaltuk azt, hogy a hangzókat nagy számban zártakként, azaz első nyíltsági fokúakként azonosították volna a hallgatók, mert azokban az esetekben, amikor az $F_1 : f_0$ hangolás megjelent, egyaránt láttunk zártabb és nyíltabb hangzókként azonosítást is. A vizsgálatunk tehát azt demonstrálta, hogy a hallgatók az operanéklésben 370 Hz alapfrekvencia fölött is azonosíthatnak második nyíltsági fokú hangzókat – szemben a beszéddel, ahol elméletileg 370 Hz alapfrekvencia felett a középső nyelvállású hangzók azonosítása (vagy azonosíthatósága) elmarad a zártabb és nyíltabb hangzókétól. Az éneklésben tapasztaltak okaként az $F_1 : f_0$ hangolás megjelenését feltételezhetjük.

A korábbi, az énekelt magánhangzók észlelését vizsgáló munkák azt találták, hogy az egyre magasabb alapfrekvencián produkciós tekintetben egyre nyíltabban ejtett beszédhangokat az észlelésben is többé-kevésbé egyre nyíltabbként, az ejtési szándék szerinti hangminőségnél magasabb F_1 -gyel rendelkező hangokként azonosítják a hallgatók (SCOTTO DI CARLO–GERMAIN 1985; BENOLKEN–SWANSON 1990; HOLLIEN et al. 2000). Bár az idézett szerzők maguk is találtak arra utaló eredményeket, hogy ezek az észlelést és az akusztikumot ilyen egyszerű, lineáris összefüggésbe hozó értelmezések nem szolgálnak a magas alapfrekvenciákon ejtett/énekelt hangzók azonosítására tökéletes magyarázattal, tanulmányaikban nem szentelnek a problémának kifejezett figyelmet. Mivel a nagyon általános tendenciákat, de főként a nagyon magas (1000 Hz körüli) alapfrekvenciákon tapasztalt nagyszámú /a/ válaszokat ez a lineáris megfeleltetés is magyarázza (vö. még a jelen fejezet vizsgálati eredményeit a második és harmadik nyíltsági fokú magánhangzókat jellemző tendenciákra), korábban nem született az énekelt magánhangzóról olyan elemzés, mely az alaphártya frekvenciaérzékenységét vagy a hallás frekvenciafelbontásának más nemlineáris (frekvenciafüggő) sajátosságait is figyelembe véve igyekezett volna magyarázni az ezzel a produkció és percepció közti lineáris megfelelés feltételezésével nem magyarázható percepció jelenségeket.

A jelen vizsgálat eredményei szerint a 370 Hz alapfrekvencia alatt és az 1000 Hz közeli alapfrekvencián énekelt magánhangzók észlelésének tendenciái jól magyarázhatók TRAUNMÜLLERnek (1981) a zártságészleletet az F_1 és f_0 tonotópiás távolságából levezető modellje segítségével. A modell keretei között értelmezve ugyanis belátható, hogyan lehetséges 370 Hz alatt a zárt, felső nyelvállású magánhangzók jó arányú felismerése – még az F_1 hangolásának, azaz az F_1 emelésének fennállása esetén is –, valamint a nyíltabb hangzók zártabbakként való felismerése, ahogy indokoltnak tűnik az is, hogy az 1000 Hz körüli alapfrekvencián énekelt hangzókat miért leginkább /a/-szerű hangoknak feleltetik meg a hallgatók. Ugyanakkor az $F_1 : f_0$ hangolás miatt úgy tűnik, hogy 370 Hz és 1000 Hz között a modell nem bír jó magyarázó erővel az énekelt magánhangzók azonosításának értelmezésében. Azt tapasztaltuk

ugyanis, hogy az $F_1 : f_0$ hangolás miatt 370 Hz fölött is megmaradnak az észleletben a középső nyelvállású hangzók – melyeknek TRAUNMÜLLER modellje szerint a 370 Hz alapprofrendencia (valamint az a fölötti alapprofrendencia-tartomány) a beszédben legalábbis nem kedvez. Ezeket, a 370 Hz fölött és 1000 Hz alatt tapasztalt percepciós tendenciákat ugyanakkor viszonylag jól magyarázza az akusztikum és percepció lineáris összefüggéseit feltételező értelmezés, mely szerint tehát az egyre nyitottabb ejtésű magánhangzókat egyre nyitottabb beszédhangokként azonosítják a hallgatók. Ezen tendenciák értelmezéséhez egy korábbi percepciós vizsgálat adatai további érdekes adalékkal szolgálnak.

Egy 1959-ben publikált kísérletében FANT a magánhangzók azonosításához szükséges minimális akusztikai információt tesztelte (valamint áttételesen az alapprofrendencia hatását is a magánhangzószzerű hangok azonosításában). A tesztben FANT nem komplex hangokat használt, a vizsgálat anyagát különböző periódusidejű, azaz különböző hangmagasságú egyszerű szinuszok képezték; a tesztben részt vevők feladata az volt, hogy az elhangzó szinuszhangokat azonosítsák valamely svéd magánhangzóként. A tesztet FANT háromszor futtatta le, minden alkalommal (körülbelül) 20 svéd anyanyelvű adatközlővel. Bár az eredmények szerint egy-egy szinuszhang több eltérő minőségű magánhangzóválaszt is kiválthatott (míg mások egy magánhangzó-minőségnek sem voltak jól megfeleltethetők), az egyes beszédhangok mégis nagyon erős összefüggéseket mutattak bizonyos frekvenciákkal. FANT összefoglalása szerint az egyes hangokra a 8. táblázatban feltüntetett frekvenciák voltak elsődlegesen jellemzőek – az adott válaszok tehát ezeken a frekvenciákon voltak a legnagyobb sűrűségűek.

8. táblázat: Az egyes magánhangzó-minőségekre jellemző frekvenciák egy egyszerű szinuszhangokkal való megfeleltetési feladat szerint (FANT 1959)

Nyíltsági fok	Magánhangzó (IPA)	Frekvencia (Hz)
1.	/u/	400
2.	/o:/	625
4.	/ɑ:/	1000
4.	/æ:/	1250
2.	/e:/	2500 (másodosorban 1600 és 3500)
1.	/i:/	4000
1.	/y:/	2000
1.	/ɤ:/	1750 (másodosorban 250)
2.	/ø:/	1750 (másodosorban 250)

Az eredményeket a magánhangzók spektrumával összevetve FANT azt találta, hogy több beszédhang, de kifejezetten az első és második nyíltsági fokú veláris magánhangzók (azaz az /u/ és /o:/) esetében a szinuszokra adott válaszok a beszédhang spektrumára emlékeztető

eloszlást mutattak, azaz a legtöbb válasz a spektrum maximumainak környékére esett. Ez az eredmény persze nem meglepő, ha belegondolunk abba, hogy a velárisok első két formánása egymáshoz igen közel helyezkedik el, sőt akár könnyen 3 Bark távolságon belülre is eshet (elsősorban a nyitottabb ejtésű velárisok esetében). Így azt mondhatjuk, hogy a velárisok esetében az első két rezonancia frekvenciasávja mindenképpen dominálja az észleletet, vagy a rezonanciák érzeti integrációja következtében létrejövő egyetlen intenzív érzeti prominencia, vagy pedig két, egymáshoz közel eső, így mindenképpen nagy intenzitású prominencia révén.

FANT (1959) vizsgálata azt mutatja, hogy a velárisok viszonylag egyértelműen leképezhetőek egyetlen (ráadásul 1000 Hz alatti) szinuszos hanghullámmal, azaz egyetlen frekvenciával, mely frekvencia tehát mind a hátul képzettséget, mind pedig a nyelvállásfokot kulcsolja (auditíve). Az eredmények – másként megragadva azt mutatják, hogy bizonyos (alap)frekvenciák bizonyos beszédhangok észleletének kedveznek. Mindezzel összhangban FANT (1959) kísérletének konkrét elrendezésétől kicsit jobban elvonatkoztatva két feltételezést is megfogalmazhatunk. Egyfelől feltehető, hogy azokban a komplex hangokban, melyekben a FANT kísérletében talált frekvenciák dominánsak, valószínűsíthető lesz egy adott nyelvállásfokú veláris magánhangzó mint magánhangzóválasz az észleletben. Másfelől pedig feltehető, hogy egy másik olyan komplex hang, amelyben emellett a domináns frekvencia mellett egy magasabb frekvenciájú második prominencia is megjelenik, viszonylag egyértelműen azonosítható lesz a velárisal megegyező nyelvállásfokú, de palatális magánhangzóként (vö. CHISTOVICH–LUBLINSKAYA 1979; CARLSON et al. 1970 eredményei a palatálisok észleléséről). Ilyen értelemben tehát azt mondhatjuk, hogy a FANT (1959) által bemutatott 1000 Hz-nél alacsonyabb frekvenciájú, kitüntetett szerepű (szinuszos) rezgések szoros összefüggést mutatnak az észlelt zártsággal.

A fejezetben bemutatott vizsgálat percepciós eredményeire pillantva feltűnő az egybeesés FANT (1959) eredményeivel – legalábbis az előbbieken megfogalmazott feltételezések értelmében. Míg a jelen vizsgálatban az első nyíltsági fokú /i: u: y:/ hangokat 400 és 600 Hz között kezdték dominánsan egy fokkal nyitottabbakként azonosítani a hallgatók, addig FANTnál – hasonlóképpen – a 625 Hz-es szinuszhang is a második nyíltsági fokú hangok megjelenését indukálta. Továbbá, bár a jelen kísérletben a második nyíltsági fokú /o: e:/ hangokat még 698 Hz alapfrekvencián is nagy arányban azonosították az ejtés szerinti nyíltsági foknak megfelelően a tesztelők, 988 Hz alapfrekvencián már radikális váltást tapasztaltunk az /a:/ válaszok megszorodása miatt, ahogyan FANTnál is azt láttuk, hogy az 1000 Hz alapfrekvenciájú szinusz a legnagyobb számban negyedik nyíltsági fokú veláris /a:/ válaszokat hívott elő.

A jelen vizsgálatban azt is tapasztaltuk, hogy a veláris hangzók némiképp eltérő viselkedést mutattak a palatálisokhoz képest, annak megfelelően, hogy az előbbieken az F_1 és F_2 közelebb, míg az utóbbiakban távolabb helyezkedik el egymástól. Azt láttuk ugyanis, hogy a veláris /u:/, melynek F_2 -je már az $F_1 : f_0$ hangolás megjelenésének alapfrekvenciájától fogva követni kezdte a $2f_0$ -t, már ettől a 349 Hz-es (illetve 494 Hz-es) alapfrekvenciától, azaz az f' - h' hangoktól kezdve nyitottabbként jelent meg az észleletben (előbb /o:/-ként majd /ɒ/-ként

és végül /a/-ként). Ezzel szemben a palatális /i: y:/ F_2 -je csak 698 Hz-en (f'') közelítette meg a $2f_0$ -t, és csak 988 Hz-en (h'') esett azzal egybe, aminek megfelelően a palatálisokban a $2f_0$ -hoz eső érzeti prominencia csak a legmagasabb alapfrekvencián volt olyan intenzív, hogy az ejtett hangzót az /a/-hoz tegye hasonlatossá. A palatálisok esetében tapasztalt spektrális eloszlás TRAUNMÜLLER (1981) modelljének megfelelően azt is jelentené, hogy a zártság észlelete megmaradt az /i: y:/ esetében a 698 Hz (f'') alatti alapfrekvenciáig (azaz kísérletünkben a h' -ig, 494 Hz-ig). Az eredmények pedig tendenciaszerűen igazolják ezt a predikciót, ugyanis azt láttuk, hogy az /i:/ és /y:/ az /u:/-nál magasabb alapfrekvenciáig tartja meg a zártság észleletét (az előbbieket, azaz a palatálisok esetében a zártság észlelete 494 Hz-ig marad fenn, míg a veláris /u:/ esetében inkább csak 394 Hz-ig).

Az itt taglalt következtetéseket szem előtt tartva megfogalmazható az a megállapítás, hogy a szopránéneklésben tapasztalható $F_1 : f_0$ hangolás szerepe a hallgató és a beszédészlelési mechanizmus szempontjából valójában kettős. Míg egyfelől a hangolás az 1000 Hz környéki alapfrekvenciákon uniformizálja a beszédhangokat (az észleleti prominencia minden magánhangzóban megegyező frekvenciája révén), másfelől egyúttal azt is lehetővé teszi, hogy a valamivel alacsonyabb (de még mindig magasnak tekinthető) alapfrekvenciákon a hangolás nélkül lehetségeshez képest többféle hangzóminőség, többféle nyíltsági fokú beszédhang is megmaradhasson az észleletben (persze úgy, hogy azok nem feltétlenül egyeznek az ejtési szándéknak megfelelő beszédhangok nyíltsági fokával). Ilyen értelemben tehát akár azok a konklúziók is megkérdőjelezhetők, melyek a magánhangzó-minőségek szempontjából egyértelműen „veszteségnek” tekintik az alapfrekvencia emelését kísérő formánshangolási stratégiát (lásd pl. JOLIVEAU et al. 2004 és GARNIER et al. 2010). Ha ugyanis nem lépne fel az $F_1 : f_0$ hangolás, akkor valószínűsíthetjük, hogy a szopránok hangtartományának javán (370 Hz fölött) csak jóval kevesebb (csak az első és negyedik nyíltsági fokba tartozó) magánhangzót volna lehetséges (nagy számban) észlelni, ami mindenképpen kisebb auditív változatosságot, és így talán kisebb esztétikai élvezetet és kevésbé a „valódi” beszédre emlékeztető szövegmondást is jelentene a szopránéneklés hallgatói számára.

8. Összegzés és következtetések

A jelen értekezésben bemutatott vizsgálatok célja a magyar magánhangzók akusztikai és percepciós jellemzőinek leírása volt a magas alapfrekvenciájú éneklésben. Kísérleteinkkel egyrészt az énekelt magánhangzók ejtését, produkcióját érintő kérdésekre kerestünk választ, másrészt pedig azt kívántuk szisztematikusan felmérni, hogy az alapfrekvencia függvényében miként dolgozódnak fel a beszédbeli megvalósulásuktól kisebb vagy nagyobb mértékben eltérő énekelt magánhangzók a beszédpercepciós rendszerben.

8.1. A hipotézisek vizsgálata

***HI:** Az énekesek a magas alapfrekvenciákon énekelt magyar magánhangzók első formánsát a megemelt f_0 értékére (vagy valamivel a fölé) hangolják akkor, ha egyébként az f_0 értéke magasabb lenne, mint a magánhangzóra a beszédben jellemző első formáns értéke. Ennek értelmében az $F_1 : f_0$ hangolás megjelenése az éneklésben a magánhangzók nyíltsági fokának (illetve beszédben jellemző F_1 -értékének) függvénye.*

Első hipotézisünket a jelen értekezésben bemutatott kísérletek egyértelműen igazolták – egyetértésben az énekelt magánhangzók akusztikai szerkezetét vizsgáló nemzetközi szakirodalom eredményeivel (vö. 1.2. *alfejezet*). Akusztikai vizsgálataink során azt találtuk, hogy a szoprán énekesek a megemelt, magas alapfrekvencia értékére emelik a magánhangzók kialakításában részt vevő első üregi rezonancia, azaz az első formáns értékét akkor, ha az alapfrekvencia megközelíti az első formáns beszédbeli megjelenésének frekvenciáját. **Az $F_1 : f_0$ hangolás az egyes magánhangzók nyíltsági fokának, illetve első formánsának függvényében jelent meg az alapfrekvencia emelésével,** a zárt, első nyíltsági fokú, felső nyelvállású /i: y: u:/ hangzóknál az f' zenei hangon (349 Hz), a nyíltabb, második nyíltsági fokú vagy középső nyelvállású /e: ø: o:/ hangzók esetében az f'' zenei hangon (698 Hz), míg a legnyíltabb, harmadik és negyedik nyíltsági fokú /ɛ ɒ a:/ esetében a h'' zenei hangon (988 Hz). Ezzel összefüggésben a legmagasabb alapfrekvencián a legnagyobb akusztikai változásokat a zárt vagy első nyíltsági fokú hangzók, míg a legkisebbeket a legnyíltabb hangzók szenvedték el (2. *fejezet*, 6. *fejezet*).

Az első formánsnak az alapfrekvenciától függő változásai mellett bizonyos mértékben a második formáns frekvenciájának szintén alapfrekvencia-függő változását is megállapíthattuk. Az alapfrekvencia emelésével ugyanis az F_2 a palatális hangzók esetében enyhe

csökkenést, míg veláris hangzók esetében erős növekedést mutatott. A változások eredményeként a vizsgálatainkban előfordult legmagasabb, 988 Hz alapfrekvenciájú, azaz az ***f''*** zenei alaphangon az összes ejteni szándékozott magánhangzó helyett egyetlen uniformizálódott F_1 - és F_2 -értékkel rendelkező magánhangzószerű hang valósult meg, melynek prominens frekvencia-összetevői az f_0 és a $2f_0$ frekvenciaértékén jelentek meg. **E hangzó minősége** (mind akusztikai, mind pedig percepciósi tekintetben) a magyar magánhangzók rendszerét tekintve **leginkább az /a:/ beszédbeli megvalósulásához hasonló** (6. és 7. fejezet).

Az első formáns megemelésének célja, hogy az énekes elkerülje azokat a helyzeteket, melyek során – a magas alapfrekvenciák éneklésekor – az alapfrekvencia magasabb lenne, mint a toldalékcső első rezonanciája. Ez ugyanis mind a hangerősségben, mind a hangszínezetben, mind pedig az adott magánhangzó minősége szempontjából nemkívánatos hatásokat eredményez: a hangerősség csökken, a hangszínezetben hirtelen változás áll be, a hangszín, azaz a magánhangzó minősége pedig megváltozik, hiszen a magánhangzóspektrum elveszíti a magánhangzó-minőség kialakítása szempontjából fontosnak tartott legintenzívebb legalsó akusztikus összetevőjét (vö. pl. SUNDBERG 1979). Bár a hangosság csökkenése megnövelt szubglottális nyomással némileg kompenzálható lenne – más szóval hangerő-növekedést idézne elő, ha az énekes a fent említett helyzetben több levegőt és nagyobb intenzitással préselne ki a tüdejéből a hangadáshoz –, a nagyobb szubglottális nyomással fenntartott, huzamosabb fonáció a hangszalagok megerőltetését okozná, tehát hosszú távon káros, és a jó technikájú éneklésben mindenképpen került megoldás. Éppen ezért az énekhang szakirodalmában mára elterjedtnek tekinthető nézet szerint az énekesek e helyett a káros technika helyett a megemelt alapfrekvencia hatására fellépő hangerőcsökkenést és hangszínváltozást a formánsok frekvenciáit befolyásoló stratégiákkal kompenzálják. Az $F_1 : f_0$ hangolás eredményeként a legalacsonyabb spektrális összetevő, azaz az alaphang vagy alapfrekvencia erősítést nyer, amely egyrészt az egész hangzóspektrum energiájának növekedését, azaz hangerő-növekedést jelent, másrészt pedig lehetővé teszi a hangszínezet jobb kontrollálását is (vö. SUNDBERG 1979, 1987). Bár az $F_1 : f_0$ hangolás e pozitív következményei teoretikusan könnyen beláthatók, az $F_1 : f_0$ hangolásra a magas alapfrekvencia okozta spektrális alul-mintavételezettség miatt kevés empirikus bizonyítékkal rendelkezünk. A spektrális alul-mintavételezettség következtében ugyanis a toldalékcső rezonanciái, azaz a formánsok nem állapíthatók meg megbízhatóan a magas alapfrekvenciájú akusztikai jel (a hangnyomás-idő függvény) Fourier-transzformáción alapuló elemzése segítségével.

Az eddigi vizsgálatok, melyek az $F_1 : f_0$ hangolás megjelenését elemezték, vagy artikulációs adatokból, vagy artikulációs adatokból akusztikai modelleken keresztül derivált akusztikai adatokból indultak ki, illetve olyan alternatív felvételi és elemzési módszertant használtak, melyek valamilyen módon lehetővé tették a toldalékcső rezonanciáinak az alapfrekvenciától független detektálását (vö. 1.2.2. *alfejezet*). A jelen értekezésben a formánsfrekvenciák meghatározásához ez utóbbi megoldást választottuk: több, korábban a beszéd- és énekhang elemzésében már sikerrel használt alternatív eszköz kombinációját használtuk a formánsfrekvenciák meghatározásához a magas alapfrekvenciájú énekelt magánhangzók-

ban. Rezonanciaadatainkat manuális inverz szűrés segítségével nyertük, melyhez a hangszalagok működését rögzítő elektrolottográfias regisztrátumokat is használtunk HENRICH és munkatársai (2004) eredményeire is alapozva HERTEGÁRD és GAUFFIN (1993), illetve SUNDBERG és munkatársai (2011, 2013) nyomán. Emellett pedig „újra-mintavételeztük” a toldalékcsovet egy, a gégénél alkalmazott külső gerjesztés, egy elektromos gége segítségével SUNDBERG (1975) módszere alapján, és az így létrejött akusztikai jelet is elemeztük. A kétféle elemzés egymást kiegészítő és validáló módon szerepelt a vizsgálatban.

H2: Mivel magas alapprofrekvenciákon az énekelt magánhangzóknak a frekvenciaszerkezete és a spektrális felbontása is megváltozik, a magas alapprofrekvencián énekelt magyar magánhangzóknak az ejtési szándék szerinti azonosítása az alapprofrekvencia emelésével törvényszerűen csökken. Az f_0 és F_1 csökkenő távolsága miatt a magas alapprofrekvenciákon ejtett magánhangzók akusztikai megjelenése az ejtési szándéktól eltérő azonosításokban inkább a zárt hangzók felismerését facilitálja.

A második hipotézis csak részben igazolódott. Egyfelől igazoltunk tekinthetjük azt a feltételezést, hogy **a magánhangzóknak az ejtési szándék szerinti azonosítási aránya csökkent az alapprofrekvencia emelésével.** Eredményeink szerint ugyanis az alapprofrekvencia emelésével egyre inkább megváltozó akusztikai szerkezetű magánhangzók egyre kevésbé voltak felismerhetők az ejtési szándék szerint, ráadásul olyan módon, hogy az akusztikai változás a legmagasabb vizsgált alapprofrekvenciáig (körülbelül 1000 Hz-ig) inkább fokozatosan, nem pedig kategorikusan változó percepciós minőségeket eredményezett. Erre abból következtethettünk, hogy az egyes ejteni szándékozott magánhangzókra az alapprofrekvencia emelésével egyre változatosabb magánhangzóválaszok érkeztek, azaz a válaszok egyre nagyobb mértékben szórtak (és egy-egy válasz egyre kisebb valószínűséggel jelent meg) a magánhangzóminőségek mentén (2., 3., 5. és 7. fejezet).

Másfelől azonban kísérleteink nem igazolták azt a feltevést, hogy a szoprán alapprofrekvencia-tartomány magasabb régióiban egyértelműen a hangzók zártabbként azonosítása volna jellemző. Eredményeink szerint **370 Hz alatt** (ami nagyjából a $f_{\#}'$ zenei hangnak felel meg) az elvárásoknak (és TRAUNMÜLLERnek a perceptuális nyíltságot az F_1 és az f_0 távolságából levezető 1981-es modellje alapján jósoltaknak) megfelelően **a nem az ejtési szándék szerinti azonosítások esetén az egyes énekelt magánhangzókat alapvetően zártabbként azonosították a hallgatók.** Ugyanakkor 370 Hz ($f_{\#}'$) fölött ez a tendencia megváltozni látszott. Egyfelől ugyanis azt tapasztaltuk, hogy **500-600 Hz alapprofrekvencia (a h' hang) környékén még átlagosan az egyes hangzók zártabbként azonosítása volt jellemző** azon ejteni szándékozott magánhangzók esetében, melyek „eredeti” (szándékolt) minőségük szerint nyíltabb és zártabb szomszédokkal is rendelkeznek. Másfelől azonban azt is láttuk, hogy a **700 Hz alapprofrekvencia (f'') magasságától felfelé már kifejezetten az egyes énekelt hangzóknak az ejtési szándéknál nyíltabbként, de különösen /ɒ/ és /a:/ hangzókként való azonosítása dominált.** Ezekről a 700 Hz alapprofrekvencia fölött, de különösen a 988 Hz alapprofrekvencián (a zenei h'' hang magasságában) tapasztalt azonosításokról azonban már azt is feltehetjük,

hogy azok nem egyszerűen a magánhangzók produkciós nyílását követő perceptuális nyílás eredményei. Az 1000 Hz alapfrekvencia körül ejtett hangzókrol ugyanis akusztikai méréseink alapján azt állapítottuk meg, hogy az azok ejtésére jellemző toldalékcso-konfiguráció első és második rezonanciája, azaz az első és második formáns éppen az f_0 és a $2f_0$ frekvencia-összetevőket erősítette mind a kilenc magyar magánhangzó esetében. Ez pedig, felnőtt női ejtésben (a kísérletben részt vevő női beszélők adatai alapján is) az /a:/ beszédbeli megvalósulására jellemző frekvenciaszerkezetet, valamint az /a/-ra jellemző prominenciaszerkezetet jelent (vö. még TRAUNMÜLLER 1981) (7. fejezet).

Megfigyeléseink, valamint az a kísérletünk, hogy TRAUNMÜLLER (1981) alapvetően a beszédejtésre kidolgozott észlelési modelljét alkalmazzuk a magas alapfrekvencián énekelt (és $F_1 : f_0$ formánshangolt) magánhangzók percepciójára, arra is következtetni engedtek, hogy az $F_1 : f_0$ hangolás feltehetőleg nem csak a homogén hangszínezet és a nagy hangerejű énekhang produkcióját segíti elő magas alapfrekvenciákon. Vizsgálataink alapján ugyanis megfogalmazhattuk azt a feltételezést is, mely szerint magasab alapfrekvenciákon szintén az $F_1 : f_0$ hangolásnak köszönhető a középső és alsó nyelvvállású, azaz második és harmadik nyíltsági fokú magánhangzók megjelenése az észleletben 370 Hz alapfrekvencia fölött; a hangolás hiányában ugyanis csak első és negyedik nyíltsági fokú hangzók megjelenését váránk. Az ily módon második és harmadik nyíltsági fokú magánhangzókként azonosított énekelt hangzók nem feltétlenül (sőt sok esetben nem) az ejtési szándék szerinti azonosítások során jelentek meg válaszként, tehát nem állíthatjuk, hogy az $F_1 : f_0$ hangolás az ejtési szándéknak megfelelő azonosítást támogatta volna. Ugyanakkor talán mégis állítható, hogy az $F_1 : f_0$ hangolásból még a hallgató észlelési mechanizmusait tekintve is hasznot hajt az énekes, hiszen a hangolás által többfajta magánhangzó-minőség jelenhet meg a percepcióban, mint az $F_1 : f_0$ hangolás nélkül, tehát az énekprodukción magasabb alapfrekvenciákon az énekelt hangzók hangszínében is változatosabb maradhat.

Eredményeink alapján egyértelműen **elutasíthatjuk azt a korábban** több tanulmány (SCOTTO DI CARLO–GERMAIN 1985; BENOLKEN–SWANSON 1990; HOLLIEN et al. 2000) által is **javasolt** – és a jelen eredmények fényében igen leegyszerűsítőnek tűnő – **feltételezést, hogy az alapfrekvencia emelésével az $F_1 : f_0$ hangolás** következtében az egyes ejteni szándékozott, de **akusztikailag egyre nyiltabbként, egyre magasabb első formánssal megvalósított magánhangzókat egyre nyiltabb hangzókként is azonosítanak a hallgatók**. A jelen vizsgálatok tanúsága szerint ugyanis a produkció és a percepció ennél sokkal komplexebb viszonya mutatkozik meg (az énekelt) magánhangzók azonosításában (is) az alapfrekvencia függvényében.

H3: Mivel magas alapfrekvenciákon az énekelt magánhangzók frekvenciaszerkezete megváltozik, továbbá a magánhangzók spektrális felbontása (a környező mássalhangzókkal alkotott szonoráns hangátmenetek spektrális felbontásával együtt) csökken, a mássalhangzó-környezetet a formánsátmenetekben rejlő akusztikai kulcsok révén nem fejt ki pozitív hatást a magánhangzóknak az ejtési szándék szerinti azonosítására. Ugyanez érvényes az izolált ejtésű

magánhangzók zöngéindítására is. Bár úgy tűnik, hogy a mássalhangzó-környezet segítheti az ejtési szándék szerinti azonosítást akkor, ha az azonosítandó magánhangzót tartalmazó hordozó hangsorok olyan értelmes szavak, melyek száma jelentősen kisebb, mint a válaszként szóba jöhető magánhangzók száma, nem találunk ilyen hatást, ha a szóba jöhető értelmes szavak száma a magánhangzók számával együtt nagy (még ha nem is feltétlenül egyenlő).

Percepció vizsgálataink – a kontextus hatását vizsgáló egyetlen korábbi tanulmány (SMITH–SCOTT 1980) eredményeivel ellentétben – azt bizonyították, hogy **sem a mássalhangzó-környezet mint a koartikulációs formánsátmenetek révén dinamikus akusztikai kulcsokat biztosító kontextus, sem pedig az izolált ejtésű magánhangzók zöngéindulása nem fejt ki pozitív hatást a magas alapprofrekvencián énekelt magánhangzóknak az ejtési szándék szerinti azonosítására (4. fejezet).** Ugyanakkor azt is kiemeltük, hogy SMITH és SCOTT (1980) valójában – szándékaiktól eltérő módon – nem a dinamikus akusztikai kulcsok, hanem az értelmes szavakba ágyazottság hatását tesztelte és mutatta ki vizsgálatában, ráadásul ezt is csak egy olyan kísérleti helyzetben, melyben a válaszként szóba jöhető értelmes szavak száma jelentősen kisebb volt, mint az izolált ejtésű magánhangzók (vagy értelmetlen szavak) száma. Emellett azt is megállapíthattuk, hogy bár a jelen kötetben vizsgált különböző képzésmódi és némiképp eltérő képzéshelyű nazális és frikatíva hangkörnyezetek valamelyest eltérő hatást mutattak a magánhangzó-azonosításra, ezek az eltérések nem voltak számottevőek, azaz **egyik vizsgált mássalhangzó-kontextus sem hatott igazán pozitívan az ejteni szándékozott magánhangzóknak az ejtési szándék szerinti azonosítására (2. és 3. fejezet).**

Míndez annak fényében nem meglepő, hogy amint tapasztaltuk, az egyes ejteni szándékozott magánhangzók megvalósulása fokozatosan változik, és az ejteni szándékozott magánhangzó-minőségtől egyre inkább eltérő lesz az alapprofrekvencia emelésével. Az akusztikus vetületük tekintetében megváltozó hangzók miatt ugyanis ésszerű feltételeznünk, hogy a magánhangzó és a szomszédos mássalhangzók közti formánsátmenetek sem kulcsolhatják az ejteni szándékozott, de valójában meg nem valósított magánhangzót – legfeljebb csak az ejtési szándéktól már többé vagy kevésbé eltérő magánhangzószerű hangot. Ennek értelmében a mássalhangzó-kontextus kizárólag és legfeljebb az ejtési szándéktól eltérő azonosítások tendenciáit egységesítheti. Arra azonban, hogy a mássalhangzó-környezet valóban szisztematikusan növelné valamely, az ejtési szándéktól eltérő hangzó azonosításának számát – azaz valamilyen módon diszkrétizálná a graduálisan változó hangzóminőségeket – nem találtunk bizonyítékot, így pusztán a jelen eredmények alapján nem állíthatjuk, hogy a mássalhangzó-környezet a formánsátmenetek révén bárminemű pozitív hatással lenne a magas alapprofrekvenciájú énekelt magánhangzók azonosítására.

Míndemellett azonban azt is láttuk, hogy a mássalhangzó-környezet mint értelmes szavakba ágyazottság nagyban befolyásolta az azonosítás tendenciáit. Bár SMITH és SCOTT (1980) célja nem ez volt, de vizsgálatukkal rávilágítottak arra, hogy a mássalhangzó-környezet pozitív hatást fejthet ki a magánhangzóknak az ejtési szándék szerinti azonosítására akkor, ha a magánhangzó és a mássalhangzók együtt értelmes szót alkotnak úgy, hogy ezzel egyúttal

a „vetélkedő megoldások” számát is csökkentik. A szerzők ugyanis a kontextus hatását egy négy értelmes szót, négy lehetséges kiegészítést, azaz négy magánhangzót mint válaszlehetőséget megengedő kondíció, valamint az angol magánhangzók rendszeréből következően legalább tízféle választ megengedő izolált magánhangzós kondíció összevetésével vizsgálták – és így mutattak ki pozitív hatást az első kondíció javára. Mindezek után már csak az maradt kérdés, hogy az értelmes szavakká kiegészíthető mássalhangzóváz akkor is pozitív hatást fejt-e ki az észlelésre, hogyha a válaszlehetőségek száma magas az észlelési feladatban, tehát a pozitív hatás kizárólag a válaszok számosságának köszönhető-e, vagy esetleg az azonosításban részt vevő kognitív folyamatokkal is összefüggésben áll. A jelen vizsgálatok egyikében éppen ezért a magánhangzók azonosítását két, kiegyenlítettebb (de még mindig nem egyenlő) számú válaszlehetőséget megengedő, értelmes, illetve értelmetlen szavakat tartalmazó kondícióban vetettük össze. Az eredmények a várakozásainknak megfelelően azt a feltevést igazolták, hogy ha megfelelően nagyszámú a válaszlehetőség mindkét kondícióban, akkor a mássalhangzós kontextus mindenképpen elveszíti pozitív hatását, a kísérletben ugyanis egyik feltétel sem segítette jobban az egyes énekelt hangzóknak az ejtési szándék szerinti felismerését, sem a jelentéssel bíró, értelmes, sem pedig a jelentés nélküli, értelmetlen szavakat tartalmazó. **A szójelentés megléte tehát, úgy látszik, önmagában nem elegendő a magánhangzóknak az ejtési szándéknak megfelelő azonosításához**, annak ellenére, hogy a vizsgált két kondícióban feltételezhetően eltérő kognitív folyamatok segítik a magánhangzók azonosítását. Ezzel együtt arra következtettünk, hogy a mássalhangzó-kontextusra korábban talált pozitív hatás (SMITH–SCOTT 1980) kizárólag annak volt köszönhető, hogy a kísérlet észlelési feladatában **a mássalhangzós kontextusban szóba jöhető értelmes szavak száma kisebb volt, mint a lehetséges izolált magánhangzók, azaz az „értelmetlen szavak” száma** (5. fejezet).

Mindezek alapján levontuk azt a következtetést is, hogy szükséges a mássalhangzó-környezet kétértelmű fogalmának pontos definíciója, az egyes tanulmányokban ugyanis sokszor keveredik a két, élesen elválasztandó aspektus: a koartikulációs formánsátmenetek és a dinamikus akusztikai kulcsok, valamint az értelmes szavakba ágyazottság.

Végezetül pedig a kontextus hatását vizsgáló kísérleteink is igazolták azt, hogy a magánhangzók akusztikai szerkezetének változása révén – a legtöbb magánhangzó esetében – valóban ellehetetlenül a magánhangzóknak az ejtési szándék szerinti azonosítása a magasabb alaphérfrekvenciákon. A kontextus hatását vizsgáló korábbi kísérletek és saját vizsgálataink alapján mint legfontosabb tanulság azt állapíthattuk meg, hogy **a magas alaphérfrekvencián énekelt, a hallgató által ismert nyelven megszólaló, értelmes szavakat tartalmazó dalokban és áriákban elhangzó magánhangzók azonosítását elsősorban nem a megvalósuló hangzók akusztikai szerkezete, hanem sokkal inkább csak statisztikai hatások befolyásolják**: az egyes, **a magánhangzót tartalmazó szavak előfordulási gyakorisága és valószínűsége** (pl. az adott kontextusban), valamint **a szavak szomszédsági sűrűsége** (*phonological neighborhood density*), azaz egy szó összetéveszthetőségének mértéke, melyet a tőle pusztán a kérdéses magánhangzóban különböző szavak száma határoz meg.

Vizsgálataink tehát azt demonstrálták, hogy a **magánhangzókat az éneklésben jellemző és az alapfrekvencia emelésével egyre fokozódó produkciós változásokat** pusztán a **beszéd-percepció folyamatok szintje nem képes kompenzálni, a kompenzációhoz a magasabb nyelvi szintekhez kötődő** (a lexikai, szemantikai, grammatikai, pragmatikai feldolgozást végző) **beszédmegértési folyamatok aktivációja szükséges**. Ezt támasztja alá az az eredményünk is, mely szerint az énekesek **produkciós gyakorlottságuk ellenére semmilyenfajta percepció előnyt nem látszanak élvezni a magas alapfrekvencián énekelt magánhangzóknak az ejtési szándék szerinti azonosításában** (5. fejezet).

8.2. A kötetben bemutatott kutatómunka fontosabb eredményeinek rövid összefoglalása

1. A magyar magánhangzókra is igazoltuk azt, a magas alapfrekvenciákon éneklő szoprán énekesekkel kapcsolatban feltételezett stratégiát, mely szerint akkor, **ha az alapfrekvencia emelése során az alapfrekvencia magasabb lenne, mint az első formáns, az énekes az első formáns frekvenciáját az alapfrekvencia értékére (vagy valamivel a fölé) emeli**. Demonstráltuk az összes magyar nem nyílt magánhangzó **produkciós nyitódását** az alapfrekvencia emelésével, valamint a **magánhangzók redukcióját és konvergálását az /a:/ pozíciója felé** az akusztikai magánhangzótérben.

2. A magas alapfrekvenciájú szopránéneklésben realizálódó magánhangzók akusztikai szerkezetének elemzésében elsőként használtuk két, már korábban alkalmazott mérési módszer kombinációját: egyfelől **manuális inverz szűréssel** elemeztük a mikrofonnal rögzített beszédjel (azaz az énekhang) hullámformáját (a mérés pontosságát elektrolottográfia rögzített oscillogram deriváltjával biztosítottuk), másfelől pedig **külső forrásjellel** (elektromos géggel) gerjesztettük a toldalékcsovet, és elemeztük az így kisugárzott hang akusztikai szerkezetét is. Ezzel az éneklés közben jellemző **toldalékcsovi rezonanciákat** (azaz a formánsfrekvenciákat) **sikerült a fonáció alapfrekvenciájától függetlenül mérnünk**.

3. Megcáfoltuk azt a szakirodalomban honoló (de nem jól kontrollált kísérletekre alapozott) feltételezést, mely szerint a magánhangzók az alapfrekvencia emelésével fokozódó produkciós nyitódását a hangzómegvalósulások fokozatos percepció nyitódása kísérné. Eredményeink szerint a produkció és percepció összefüggése e kérdésben (is) jóval komplexebb. **Az alapfrekvencia emelésével alapvetően a magánhangzók percepció minőségének egyre fokozódó elbizonytalanodását tapasztaltuk**. A legzártabb hangzók észlelt minősége már 3-400 Hz körül, azaz az *f'-h'* hangok magasságában bizonytalanává válik, míg a nyíltabb hangzók (az /*ɒ a:/* kivételével) 700 Hz azaz az *f'''* környékén „veszítik el” szándékozott minőségüket. A tendenciának leginkább ellenálló legnyíltabb magánhangzók, az /*ɒ a:/* azonosítása nem változik szisztematikusan az alapfrekvenciával. Az énekelt

magánhangzók észlelésére vonatkozó tendenciákat részletesebben a következő két tézisben összegezhetjük.

3. a) A produkciós adatoknak megfelelő, de **az ejtési szándéknál nyíltabb (/b a:/ hangzóként történt) magánhangzó-azonosításokat** csak abban az esetben **találtunk**, amikor az alapfrekvencia kellően magas volt ahhoz, hogy az f_0 és a $2f_0$ egymáshoz viszonyított távolsága, valamint a két összetevő erősítettsége is hasonlóvá tegye az egyes magas alapfrekvencián énekelt magánhangzószerű hangzókat az /a:/ beszédbeli prominenciamintázatához. Ez az $F_1 : f_0$ hangolás miatt csak akkor valósul meg egyértelműen, amikor az alapfrekvencia emelésével változó frekvenciaértékű F_2 formáns eléri (és erősíteni kezdi) a $2f_0$ összetevőt, azaz megközelítőleg **a zenei h" hang (988 Hz alapfrekvencia) magasságában**.

3. b) Az 1000 Hz alatti alapfrekvenciákon a produkció és percepció az egyszerű lineáris összefüggésnél sokkal komplexebbnek mutatkozott: **5-600 Hz-ig** még alapvetően **az egyes „eredeti”** minőségük szerint nyíltabbként és zártabbként is azonosítható **hangzók az ejtési szándéknál zártabbként azonosítása dominált, az ejtési szándéknál nyíltabb azonosítások száma csak e fölött az alapfrekvencia-sáv fölött kezdett el növekedni**. Eredményeink alapján feltételezhetjük, hogy az egyes, az $F_1 : f_0$ formánshangolás következtében produkciósan nyíltabb énekelt magánhangzók az ejtési szándéknál nyíltabbként történő azonosításához szükséges a spektrális alul-mintavételezettség fokozódása, és az F_2 és $2f_0$ egybeesése (azaz a $2f_0$ erősítése) is, azaz megállapíthatjuk, hogy **a percepció nyitódás (azaz az egyes megvalósított magánhangzók nyíltabb magánhangzóként való felismerése) nem a produkcióban tapasztalt nyitódás egyenes következménye**.

4. Valószínűsítjük, hogy a magánhangzók akusztikai szerkezetét jelentősen megváltoztató **$F_1 : f_0$ hangolás következtében lehetséges az, hogy második és harmadik nyíltsági fokú magánhangzók is megjelennek a percepcióban magas alapfrekvenciákon** (bár nem feltétlenül az ejtési szándéknak megfelelően) – a formánshangolási tendencia nélkül elsősorban ugyanis csak az első és negyedik nyíltsági fokú hangzók megjelenését jósolná a szakirodalom magas f_0 esetén.

5. Megállapítottuk, hogy **a mássalhangzó-környezet a dinamikus akusztikus kulcsok, azaz a koartikulációs formánsátmenetek révén nem segíti a magas alapfrekvencián énekelt magánhangzóknak az ejtési szándék szerinti azonosítását**. Ugyanakkor azt is megállapíthatjuk, hogy **a mássalhangzó-környezet mint értelmes szavakba ágyazottság pozitív hatást fejthet ki az ejtési szándék szerinti azonosításra** akkor, ha az adott szó, illetve magánhangzó valószínűsége magas, tehát a szónak kevés fonológiai szomszédja van (ti. csak kisszámú olyan szószomszédal rendelkezik, melytől csak a kérdéses magánhangzó különbözteti meg).

6. Megállapítottuk, hogy az énekes és nem énekes (tehát énektechnikailag képzetlen) hallgatóság gyakorlatilag ugyanolyan tendenciák szerint azonosítja a magas alapfrekvencián énekelt magánhangzókat, tehát **az éneklésben szerzett gyakorlat nem jelent percepció előnyt a magánhangzóknak az ejtési szándék szerinti felismerésében**.

7. Megállapítottuk, hogy **az énekelt magánhangzók akusztikai szerkezete az alapfrekvencia emelésével egyre jobban megváltozik** (azaz egyre jobban eltér a magánhangzók

beszédbeli megvalósulásától), mely produkciós tendencia pusztán a beszédpercepció szintjén nem kompenzálható. A kompenzációhoz, tehát **az egyes énekelt szavak azonosításához/felismeréséhez a magasabb szintű beszédmegértési folyamatok aktivációja szükséges.**

8.3. A vizsgálatok közvetlen és közvetett hozadékai, kitekintés

A valóságban az alapszófajú, több szótagos szavak szomszédsági sűrűsége (*phonological neighborhood density*) alacsonyabb, mint a jelen kísérletekben vizsgált szavaké, valamint alapvetően nem jellemző az értelmetlen szavak ejtése a klasszikus operaéneklés során. Ugyanakkor az is igaz, hogy a szavakat általánosságban erősen szótagolva ejtik az énekesek, ráadásul úgy, hogy a szótagokból több is a kritikus alapfrekvenciákon felül valósul meg. Így az **értelmetlen szótagokban vagy izolációban ejtett magánhangzók azonosítását vizsgáló kísérleteknek** nem csak azért van relevanciája és jelentősége, mert ezekkel **differenciáltan a beszédészlelés folyamatát vizsgáljuk, és a beszédészlelés a beszédmegértés alapját jelenti** (tehát nincs beszédmegértés beszédészlelés nélkül), hanem azért is, mert **valószínűleg nem ritka, hogy a valóságban is az itt látott**, mesterségesen előállított körülményekhez nagyon hasonló **körülmények között kell azonosítanunk egy-egy szótagot az operaéneklés hallgatása során.** Éppen ezért az itt tárgyaltak függvényében feltehetőleg jó és támogatandó gyakorlatnak tekinthető – bár sokak számára talán kissé furcsának hathat –, az operaszövegek feliratozása, valamint a dalesteken, koncerteken a hallgatóság számára valamilyen formában prezentált szövegeknyv még a beszélők anyanyelvén megszólaló előadások esetén is.

Ezen túlmenően a magánhangzók alapfrekvencia-függő megvalósítása feltehetőleg szem előtt tartandó **szempont lehet(ne) az operaszövegek fordításában** is. Míg egyes kutatások szerint például Wagner kifejezetten ügyelt rá, hogy a magasabb alapfrekvenciákon harmadik és negyedik nyíltsági fokú (jellemzően /a/) hangzókkal „könnyítse meg az énekes dolgát” (SMITH–WOLFE 2009), egy rosszul sikerült fordítás könnyedén semmissé teheti ezt a törekvést, és megnehezítheti mind az énekes produkciós, mind pedig a hallgató percepció feladatát.

Elgondolkodásra érdemes kérdés, hogy **eredményeink** mennyiben vehetők figyelembe, vagy netán mennyire alkalmazhatók (ha alkalmazhatók egyáltalán) **az énektanítás módszertanában.** Tudjuk, hogy sokszor bizonyos alapvetően tudatosan „nehezen elérhető” artikulációs mozgások – például a gége függőleges helyzetének direkt változtatása – csak akkor hajthatók végre hatékonyan, ha az ejtéshez kiadott tudatos motoros parancsok nem közvetlenül az adott artikulációs szerv működését igyekeznek befolyásolni.³³ Ezzel összefüggésben

³³ Ezt használják ki például a fül-orr-gégészek mindennapos praxisukban, mikor a gégetükrözéshez az /i/ hangzó ejtésére kéri a páciens úgy, hogy közben leszorítják a nyelvet, és szétfeszítik az ajakat, tehát valójában teljesen ellehetetlenítik az /i/ ejtését. Azért kéri mégis, hogy ezt a hangzót próbálja ejteni a vizsgált személy, mert

például úgy is okoskodhatunk, hogy feltesszük, hogy a különböző, az énekhang minősége szempontjából fontos produkciós jellemzők, akusztikus hatások eléréséhez is szükséges lehet az az illúzió, mely szerint az egyes magánhangzók még magas alapfrekvenciákon is elkülöníthetőek a produkcióban. Amennyiben ez így van, nem lehetünk biztosak benne, hogy egy már kiforrott és begyakorolt énektechnikával dolgozó énekes számára hasznos lehet az az információ, hogy a beszédhangok magas alapfrekvenciákon valójában uniformizálódnak, hiszen ő már nagy rutinnal (azaz jobbára öntudatlanul, automatizált technikával) és vélhetőleg hatékonyan működteti az artikulációs szerveit az éneklés során. Ugyanakkor egy még tanuló, kezdő énekes számára talán segítség lehet, ha a megfelelő hangszín kialakításában nem a magánhangzó-minőségek fenntartását és elkülönítését hangsúlyozza, hanem más akusztikus és esztétikai szempontokat érvényesít az oktatás. Egyúttal egy szintén ehhez kapcsolódó, és cseppet sem elhanyagolható aspektus az sem, hogy az énekeseket láthatólag megnyugtatja egy, a magas alapfrekvencián énekelt beszédhangok elkülönítésének ellehetetlenüléséről szóló tudományos bizonyíték. Ezt bizonyították a szerző személyes beszélgetései is énekes adatközlőivel, de jól példázza ezt a következő idézet is egy énekes (nem akusztikai fonetika, sokkal inkább előadóművészeti irányultságú) disszertációjából: „a tanult technikák segítségével biztonsággal létrehozhatjuk azt a hangot, ami megfelelő dinamikával hallható a nagyzenekari hangzás mellett, de rendkívül magas hangtartományban és a hosszan tartott hangzás esetén könnyen szövegtorzulást eredményez. Ez nemcsak a zene és a prozódia (esetleges fordítás) szerkesztési problémája, hanem teljesen normális akusztikai folyamat eredménye” (Váradi 2010: 36).

Végezetül pedig érdemes kitérnünk általánosságban is **az akusztikai és artikulációs fonetika** némenklatúrájának, eszközeinek, vívmányainak alkalmazhatóságára, valamint az alkalmazás létjogosultságára **az énektanítás módszertanában**. Vélekedésünk szerint az énekprodukciónról szóló szakmai diskurzus, illetve az éneklés mint (beszéd)produkciós mód tanítása feltehetőleg sokat profitálhat abból, ha a hangképzés eszközeit (az artikulációs szerveket) és eredményét (az akusztikumot) minél objektívebb eszközökkel kezeli, és azokra a lehető legobjektívebb megnevezésekkel referál – mindezt pedig a fonetika tudománya teszi lehetővé. Ez a felvetés ugyanakkor nem csak „kívülről” látszó, hanem – a tapasztalatok szerint – az énekes szakma berkein belül is felmerülő igény. Az énektechnika tanításával kapcsolatban egy manapság meglehetősen gyakran felmerülő kritika például éppen azt nehezményezi, hogy a tanárok az egyes produkciós, akusztikus jellegek megragadására általában nehezen „lefordítható”, nem mindenki számára jól érthető, absztrakt metaforákat (pl. „előlképzett” hang, „nyitott” hang stb.) használnak az oktatásban, valamely sokkal kevésbé relatív eszköz- és fogalomtár helyett.

köztudottan az /i/ ejtésekor áll minden beszédhang közül a legmagasabban a gége, tehát a vizsgálat során az /i/ ejtésére törekedve a páciens valójában (öntudatlanul) a lehető legmagasabb, a gégetükör számára is látható helyzetbe emeli a gégét. Közvetlenül a gége emelésére azonban valószínűleg hiába kérné az orvos a beteget, hiszen a gége helyzetét a legtöbben nem vagyunk képesek közvetlenül befolyásolni.

A metaforák használata – szemben a sokkal egyértelműbb anatómiai és akusztikai fogalmakkal – például az eltérő fiziológiájú énekesek esetében komoly problémákat okozhat. Ahhoz, hogy megértsük, ez miért is van így, azt kell látnunk, hogy az anatómiai-akusztikai fogalmak használata fejleszti az artikulációs-akusztikai tudatosságot is, tehát segít annak megértésében, hogy mely képzőszervekkel mely hangjellemzőket tudjuk befolyásolni és miként. Ha azonban az objektív terminusok helyett metaforákat használunk, ez a fajta tudatosság nem fejlődik, és az adott hangszín eléréséhez is csak a próbálgatás és a tanárnak az ejtést érintő, de kizárólag az önmaga megfigyelésén alapuló és leginkább intuitív-metaforikus artikulációs „leírásának” imitációja marad. Noha ez a tanítási stratégia hasonló fizikumú tanárok és diákok találkozásakor eredményes lehet, eltérő testi (elsősorban toldalékcsovi) adottságok mellett általában nem, hiszen más fiziológiai adottságok mellett más artikulációs-akusztikai eszközök szükségesek ugyanazon hangélmény kialakításához. Mindez tehát nem jelent mást, mint azt, hogy ha egy eltérő fiziológiai adottságokkal rendelkező tanár-diák páros nem az egyértelmű, objektív megnevezéseket használja, és nem törekszik a fiziológiai-akusztikai működések tudatosítására, nem is igen lehet képes jól megérteni egymást, hiszen a tanár által alkalmazott artikulációs stratégiák a más fiziológiai adottságokkal rendelkező diák esetében nagy valószínűséggel nem érik el a kívánt akusztikai hatást, hangszínezetet.

Énekes kísérleti személyeimtől szerzett értesüléseim szerint valóban nem ritka, hogy egy énektanár és diákja közti munkakapcsolat a kettejük közti fiziológiai különbségek miatt hiúsuljon meg, melynek oka tehát nem más, mint az általuk használt metaforikus nyelvezet és annak (negatív) hatása az artikulációs-akusztikus tudatosság fejlődésére. Ugyanakkor a beszédképző szervek anatómiájára és fiziológiájára támaszkodó oktatásnak, tehát annak a módszertannak, mely a tanár és diákja közti fiziológiai különbségekből fakadó problémákat a fizikai valóságot megragadó instrukciók révén áthidalni képes, már van hagyománya. Nagyban ezen alapszik például az iskolateremtő Jo Estill vagy Cornelius Reid oktatásmódszertana is.

Zárásképp a szerző kénytelen felhívni a figyelmet könyve egy nagy hiányosságára – adósok maradtunk ugyanis azzal, hogy megvizsgáljuk és ismertessük a magyar köznyelvi magánhangzóknak a szopránéneklésben tapasztalható artikulációs változásait. Ehhez kapcsolódóan hangsúlyoznunk kell, hogy az artikulációs vizsgálatok eszköztára (lásd pl. a 1.2.1. és 1.2.2. *alfejezeteket*) korántsem olyan könnyen (és olcsón) hozzáférhető, mint az akusztikai szerkezet vagy az észlelési folyamatok elemzéséhez használt technikák. Ez hazánkban ráadásul olyannyira is igaz, hogy az artikulációs vizsgálatokhoz használható eszközök jó része – legalábbis a közelmúltig – még egyáltalán nem volt elérhető beszédvizsgálatokra. Éppen ezért Magyarországon az utóbbi évtizedekben csak elvétve végeztek artikulációs fonetikai vizsgálatokat. Mindennek ellenére a jelen kötet szerzője kutatótársaival abban a szerencsében részesült, hogy egy német–magyar együttműködési projekt keretében megkezdhette a magyar köznyelvi magánhangzókra az éneklésben jellemző artikulációs változások vizsgálatát.³⁴

³⁴ A magyar kutatócsoport vezetője dr. Markó Alexandra, tagjai a szerzőn kívül Bartók Márton, Jankovics Julianna és Weidl Zsófia, akik egyetemi hallgatóként kapcsolódtak be a kutatócsoport munkájába. A német

A kutatómunkából, mely a nyelv és az állkapocs mint artikulációs szervek működését elemzi elektromágneses artikulográfiával magyar és német anyanyelvű énekesek ejtésében, már egy szoprán énekesnő (az első kísérleti személy) ejtésére vonatkozóan megszületett az első eset-tanulmány-jellegű elemzés és publikáció (DEME et al. 2016).

Szándékunk szerint az éneklésben megvalósuló beszédhangok vizsgálatában a jövőben tehát ezt az irányt, azaz az énekelt beszédhangok artikulációs vizsgálatát kívánjuk követni mind a magánhangzó-, mind pedig a mássalhangzóejtés tekintetében. Az artikulációs működések további vizsgálata remélhetőleg segítségünkre lesz abban, hogy pontosabb képet alkothassunk a kötetben ismertetett akusztikus jellemzők kialakításáért felelős artikulációs stratégiákról az egyes magánhangzó-minőségek tükrében, így várhatóan sikerülhet tisztáz-nunk néhány, a nemzetközi szakirodalomban felvetett, de eddig még kielégítő válasz nélkül hagyott kérdést is.

kutatócsoport vezetője prof. dr. Reinhold Greisbach, tagja Michelle Meier mesterszakos egyetemi hallgató. A kutatást a TKA–DAAD projektalapú kutatócsere program támogatja (64685 sz.).

9. Irodalom

- American Standards Association 1960. *American Standard Acoustical Terminology*. American Standards Association, New York.
- ASHBY, Patricia 2011. *Understanding phonetics*. Hodder Education, London.
- BEDDOR, Patrice S. 1983. *Phonological and phonetic effects of nasalization on vowel height*. PhD thesis. Indiana University Linguistics Club, Bloomington.
- VON BÉKÉSY, Georg 1960. *Experiments in hearing*. McGraw-Hill, New York.
- BENADE, Arthur H. 1976. *Fundamentals of musical acoustics*. Oxford University Press, London.
- BENOLKEN, Martha S. – SWANSON, Charles E. 1990. The effect of pitch-related changes on the perception of sung vowels. *Journal of the Acoustical Society of America* 87/4. 1781–1785.
- BERGER, Kenneth W. 1964. Some factors in the recognition of timbre. *Journal of the Acoustical Society of America* 26/10. 1888–1891.
- BOERSMA, Paul – WEENINK, David 2009. *Praat: Doing phonetics by computer (Version 5.3)*. <http://www.praat.org>. (A letöltés ideje 2014. január 12.)
- BOLLA Kálmán 1995. *Magyar fonetikai atlasz. A szegmentális hangszerkezet elemei*. Nemzeti Tankönyvkiadó, Budapest.
- BÓNA Judit 2012. A zöngétlen explozívák zöngékezdési ideje idők és fiatalok felolvasásában. In GÓSY Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest, 121–137.
- BÓNA Judit 2014. Temporal characteristics of speech: The effect of age and speech style. *Journal of the Acoustical Society of America* 136/2. EL116–EL121.
- BÓNA Judit – IMRE Angéla 2010. A rövid-hosszú magánhangzók óvodás és kisiskolás gyermekek beszédprodukcójában. In NAVRACSICS Judit (szerk.): *Nyelv, beszéd, írás. Pszicholingvisztikai tanulmányok I*. Tinta Könyvkiadó, Budapest, 49–56.
- BRANDERUD, Peter – McALLISTER, Robert – KASSLING, Bo 1994. Methodological studies of Motetrack: Coil tilt and placement. *PERILUS XVIII*. Stockholm University, Stockholm, 91–109.
- BRESCH, Erik – NARAYANAN, Shrikanth 2010. Real-time magnetic resonance imaging investigation of resonance tuning in soprano singing. *Journal of the Acoustical Society of America* 128/5. EL335–EL341.
- CARLSON, Rolf – FANT, Gunnar – GRANSTRÖM, Björn 1975. Two-formant models, pitch and vowel perception. In FANT, Gunnar – TATHAM, M. A. A. (eds.): *Auditory analysis and perception of speech*, Academic Press, London, 55–82.
- CARLSON, Rolf – GRANSTRÖM, Björn – FANT, Gunnar 1970. Some studies concerning perception of isolated vowels. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 2–3. 19–35.

- DE CHEVEIGNÈ, Alain – KAWAHARA, Hideki 1999. Missing-data model of vowel identification. *Journal of the Acoustical Society of America* 105/6. 3497–3508.
- CHIBA, Tsutomu – KAJIYAMA, Masato 1941. *The vowel: Its nature and structure*. Tokyo: Kaiseikan, Tokyo.
- CHISTOVICH, Ludmilla A. – LUBLINSKAYA, Valentina V. 1979. The ‘center of gravity’ effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli. *Hearing Research* 1/3. 185–195.
- COOKER, Harry S. 1973. On the problem of tracking mandibular movements. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 14/2–3. 1–012.
- DELGUTTE, Bertrand – KIANG, Nelson Y. S. 1984. Speech coding in the auditory nerve: I. Vowel-like sounds. *Journal of the Acoustical Society of America* 75/3. 866–878.
- DEME, Andrea – GRÁCSI, Tekla Etelka – JANKOVICS, Julianna 2013. Obstruent voicing in singing. Presentation. *15th Summer School of Psycholinguistics*. Balatonalmádi, 2013 May 26–30.
- DEME, Andrea – GREISBACH, Reinhold – MARKÓ, Alexandra – MEIER, Michelle – BARTÓK, Márton – JANKOVICS, Julianna – WEIDL, Zsófia 2016. Tongue and jaw movements in high-pitched soprano singing: A case study. *Beszédkutató* 24. 121–138.
- EPPS, Julien – SMITH, John R. – WOLFE, Joe 1997. A novel instrument to measure acoustic resonances of the vocal tract during speech. *Measurement Science and Technology* 8. 1112–1121.
- FANT, Gunnar 1959. *Acoustic analysis and synthesis of speech with applications to Swedish*. Ericsson Technics No. 1.
- FANT, Gunnar 1960. *Acoustic theory of speech production*. Mouton & Co., The Hague.
- FLEMMING, Edward 2005. Linguistic Phonetics, Fall 2005. Course Number 24.963. Massachusetts Institute of Technology: MIT OpenCourseWare. <http://ocw.mit.edu>. License: Creative Commons BY-NC-SA. (A letöltés ideje: 2015. március 15.)
- GARCIA, Manuel 1911. *Garcia’s treatise on the art of singing*. Leonard & Co, London.
- GARNIER, Maëva – HENRICH, Natalie – SMITH, John – WOLFE, Joe 2010. Vocal tract adjustments in the high soprano range. *Journal of the Acoustical Society of America* 127/6. 3771–3780.
- GÓSY Mária 1989. *Beszédészlelés*. Linguistica, Series A, Studia et Dissertationes 2. MTA Nyelvtudományi Intézet, Budapest.
- GÓSY Mária 1999. Az egyéni hangszínezet és a beszélő felismerésének kísérleti-fonetikai megközelítése. *Magyar Nyelvőr* 123. 424–438. <http://www.c3.hu/~nyelvor/period/1234/123406.htm>. (A letöltés ideje: 2015. február 5.)
- GÓSY Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- GÓSY Mária 2012. Az alsóbb nyelvvállású magyar magánhangzók formánszerkezete. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest, 43–66.
- GÓSY, Mária 2013. Inter-speaker and intra-speaker variability indicating a synchronous speech sound change. In SZIGETVÁRI, Péter (ed.): *VLLxx: Papers in linguistics presented to László Varga on his 70th birthday*. Tinta Könyvkiadó, Budapest, 313–332.

- GÓSY Mária – BÓNA Judit 2014. Magánhangzók ejtése fiatalok és idősek spontán beszédében. *Magyar Nyelv* 110. 129–143.
- GOTTFRIED, Terry L. – CHEW, Stephen L. 1986. Intelligibility of vowels sung by a countertenor. *Journal of the Acoustical Society of America* 79/1. 124–130.
- GRÁCZI Tekla Etelka – HORVÁTH Viktóra 2010. A magánhangzók realizációja spontán beszédben. *Beszédkutatás* 2010. 5–16.
- GRAMMING, Patricia – NORD, Lennart – SUNDBERG, Johan – ELIOT, Ninni 1993. Does the nose resonate during singing? *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 34/4. 35–42.
- GRANQVIST, Svante é. n. *DeCap*. www.tolvan.com. (A letöltés ideje: 2013. április 26.)
- HENRICH, Nathalie – D’ALESSANDRO, Christophe – DOVAL, Boris – CASTELLENGO, Michèle 2004. On use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *Journal of the Acoustical Society of America* 115/3. 1321–1332.
- HERBST, Christian T. – FITCH, W. Tecumseh S. – SVEC, Jan G. 2010. Electroglottographic wavegrams: A technique for visualizing vocal fold dynamics noninvasively. *The Journal of the Acoustical Society of America* 128/5. 3070–3078.
- HERTEGÅRD, S. – GAUFFIN, J. 1993. Voice source – vocal tract interaction during high-pitched female singing. In: *Proceedings of the Stockholm Music Acoustics Conference 1993 (SMAC 1993)*. Royal Swedish Academy of Music, Stockholm, 177–182.
- HICKOK, Gregory 2010. The role of mirror neurons in speech perception and action word semantics. *Language and Cognitive Processes* 25/6. 1–28.
- HOLLIEN, Harry – MENDES-SCHWARTZ, Ana P. – NIELSEN, Kenneth 2000. Perceptual confusions of high-pitched sung vowels. *Journal of Voice* 14/2. 287–298.
- HONDA, Kiyoshi 1983. Relationship between pitch control and vowel articulation. In STUDDERT-KENNEDY, Michael – O’BRIAN, Nancy (eds.): *Status Report on Speech Research: A Report on the Status and Progress of Studies on the Nature of Speech, Instrumentation for its Investigation, and Practical Applications*. Haskins Laboratories, New Haven, 269–282.
- HONDA, Kyioshi é. n. *Control of hyoid-larynx complex by Jo Estill*. Presentation.
- HOOLE, Philip – KROOS, Christian 1998. Control of larynx height in vowel production. In: *Proceedings of the 5th Conference on Language Processing (ICSLP)* 2. 531–534.
- HORVÁTH Viktória 2005. A magánhangzók nazalizációjáról. *Beszédkutatás* 2005. 51–62.
- HORVÁTH Viktória 2008. Nazális hatás a magánhangzóképzés függvényében. *Beszédkutatás* 2008. 22–32.
- HURME, Pertti – SONNINEN, Aato 1995. Vertical and saggital position of the larynx in singing. In ELENIUS, Kjell – BRANDERUD, Peter (eds.): *Proceedings of the XIII International Congress of Phonetic Sciences*. Stockholm, 214–217.
- International Phonetic Association 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University, Cambridge.

- International Phonetic Association 2015. *The International Phonetic Alphabet* (revised to 2015). [https://www.internationalphoneticassociation.org/sites/default/files/IPA_chart_\(C\)2005.pdf](https://www.internationalphoneticassociation.org/sites/default/files/IPA_chart_(C)2005.pdf) (A letöltés ideje: 2016. július 15.)
- JENKINS, James J. – STRANGE, Winifred – EDMAN, Thomas R. 1983. Identification of vowels in "vowelless" syllables. *Perception & Psychophysics* 34/5. 441–450.
- JOHANSSON, Christine – SUNDBERG, Johan – WILBRAND, Hermann 1982. X-ray study of articulation and formant frequencies in two female singers. In ASKENFELT, Anders – FELICETTI, S. – JANSSON, Erik – SUNDBERG, Johan (eds.): *Proceedings of the Stockholm Music Acoustics Conference 1983 (SMAC 1983)*. Royal Swedish Academy of Music, Stockholm, 203–218.
- JOLIVEAU, Elodie – SMITH, John – WOLFE, Joe 2004. Vocal tract resonances in singing: The soprano voice. *Journal of the Acoustical Society of America* 116/4. 2434–2439.
- JONES, Daniel 1922. *An outline of English phonetics*. G. E. Stechert & Co., New York.
- KÁLLAI János – BENDE István – KARÁDI Kázmér – RACSMÁNY Mihály 2008. *Bevezetés a neuropszichológiába*. Medicina, Budapest.
- KASSAI Ilona 1998. *Fonetika*. Nemzeti Tankönyvkiadó, Budapest.
- KERÉNYI Miklós György 1959. *Az éneklés művészete és pedagógiája*. Zeneműkiadó, Budapest.
- KOVÁCS Magdolna 1998. A spektrális minőség és az időtartam szerepe a magánhangzók percepciójában. *Beszédkutatás 1998*. 35–46.
- KRAKOW, Rena A. – BEDDOR, Patrice S. – GOLDSTEIN, Louis M. – FOWLER, Carol A. 1987. Coarticulatory influences on the perceived height of nasal vowels. *Status Report on Speech Research*. Haskins Laboratories, 31–53.
- LADEFOGED, Peter 1967. *Three areas of experimental phonetics*. Oxford University Press, London, 62–75.
- LADEFOGED, Peter – BROADBENT, Donald E. 1957. Information conveyed by vowels. *Journal of the Acoustical Society of America* 29/1. 98–104.
- LEE, Sungbok – POTAMIANOS, Alexandros – NARAYANAN, Shrikanth 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America* 105/3. 1455–1468.
- LIBERMAN, Alvin M. – MATTINGLY, Ignatius G. 1985. The motor theory of speech perception revised. *Cognition* 21/1. 1–36.
- LINDBLOM, Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In HARDCASTLE, William – MARCHAL, Alain (eds.): *Speech production and speech modeling*. Kluwer, Dordrecht, 403–439.
- LINDBLOM, Björn – BROWNLEE, S. – DAVIS, Barbara – MOON, Seung-Jae 1992. Speech transforms. In: *ESCA Workshop on phonetics and phonology of speaking styles: Reduction and elaboration in speech communication (PPoSpSt-1991)*. Barcelona, 357–368. http://www.isca-speech.org/archive_open/ppospst. (A letöltés ideje: 2014. március 16.)
- LINDBLOM, Björn E. – SUNDBERG, Johan 1971. Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustical Society of America* 50/4. 1166–1179.

- MÁDY Katalin 2008. Magyar magánhangzók vizsgálata elektromágneses artikulográffal gyors és lassú beszédben. *Beszédkutató* 2008. 52–66.
- MÁDY Katalin 2013. A beszédpercepció helye a teljes megértési folyamatban. In PLÉH Csaba (szerk.): *Általános Nyelvészeti Tanulmányok XXV: A kognitív szempont a nyelv pszichológiájában*. Akadémiai Kiadó, Budapest, 103–141.
- MÁDY, Katalin – REICHEL, Uwe D. 2007. Quantity distinction in the Hungarian vowel system – just theory or also reality? In TROUVAIN, Jürgen – BARRY, William J. (eds.): *Proceedings of the 16th International Congress of Phonetic Sciences*. Pirrot GmbH., Saarbrücken, 1053–1056.
- MARKÓ Alexandra 2013. *Az irreguláris zöngé funkciói a magyar beszédben*. ELTE Eötvös Kiadó, Budapest.
- MARKÓ, Alexandra – GRÁCZI, Tekla Etelka – BÓNA, Judit 2010. The realization of voicing assimilation rules in Hungarian spontaneous and read speech: Case studies. *Acta Linguistica Hungarica* 57/2–3. 210–238.
- McKINNEY, James 1994. *The diagnosis and correction of vocal faults*. Genovex Music Group, Springdale.
- MITCHELL, Helen 2005. *Defining vocal quality in female classical singers: Pedagogical, acoustical and perceptual studies*. PhD thesis. University of Sydney, Sydney.
- NEARY, Terence M. 1989. Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America* 85/5. 2088–2113.
- NÉMETH Géza – OLASZY Gábor (szerk.) 2010. *A magyar beszéd. Beszédkutató, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest.
- PABST, Friedemann – SUNDBERG, Johan 1992. Tracking multi-channel electroglottograph measurement of larynx height in singers. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 33/2–3. 67–78.
- PETERSON, Gordon E. – BARNEY, Harold L. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24/2. 175–184.
- R Core Team 2013. *R: A Language and Environment for Statistical Computing*. Vienna. <http://www.R-project.org>.
- ROGERS, Catherine L. – DEMASI, Teresa M. – KRAUSE, Jean C. 2010. Conversational and clear speech intelligibility of /bVd/ syllables produced by native and non-native English speakers. *Journal of the Acoustical Society of America* 128/1. 410–423.
- ROSNER, B. S. – PICKERING, J. B. 1994. *Vowel perception and production*. Oxford University Press, Oxford.
- ROTHENBERG, Martin 1973. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *The Journal of the Acoustic Society of America* 53/6. 1632–1645.
- ROUBEAU, Bernard – HENRICH, Nathalie – CASTELLENGO, Michèle 2009. Laryngeal vibratory mechanisms: The notion of vocal register revisited. *Journal of Voice* 23/4. 425–438.
- RYALLS, John H. – LIBERMANN, Philip 1982. Fundamental frequency and vowel perception. *Journal of the Acoustical Society of America* 72/5. 1631–1634.

- SALDANHA, E. L. – CORSO, John F. 1964. Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America* 36/11. 2021–2026.
- SCOTTO DI CARLO, Nicole – GERMAIN, Aline 1985. A perceptual study of the influence of pitch on the intelligibility of sung vowels. *Phonetica* 42/2. 188–197.
- SEIKEL, Anthony J. – KING, Douglas, W. – DRUMRIGHT, David G. 2010. *Anatomy & Pshysiology for Speech, Language and Hearing*. 4th edition. Cengage Learning, Delmar.
- SIPTÁR Péter 2014. Fonológiai ábrázolás és fonetikai megvalósítás: a magyar /a:/ esete. *Magyar Nyelv* 110. 335–342.
- SIPTÁR, Péter – TÖRKENCZY, Miklós 2000. *The phonology of Hungarian*. Oxford University Press, Oxford.
- SJÖLANDER, Kåre – BESKOV, Jonas 2000. Wavesurfer – an open source speech tool. In YUAN, B. – HUANG, T. – TANG, X. (eds.): *Proceedings of ICSLP 2000, 6th International Conference on Spoken Language Processing*. Beijing, 464–467. [http://: www.speech.kth.se/wavesurfer](http://www.speech.kth.se/wavesurfer) (A letöltés ideje: 2013. február 14.)
- SMITH, John – WOLFE, Joe 2009. Vowel-pitch matching in Wagner’s operas: Implications for the intelligibility and ease of singing. *Journal of the Acoustical Society of America* 125/5. 196–201.
- SMITH, Lloyd – SCOTT, Brian L. 1980. Increasing the intelligibility of sung vowels. *Journal of the Acoustical Society of America* 67/5. 1795–1797.
- SRULOVICZ, Peter – GOLDSTEIN, Julius L. 1983. A central spectrum model: A synthesis of auditory-nerve timing and place cues in nonaural communication of frequency spectrum. *Journal of the Acoustical Society of America* 73/4. 1266–1276.
- Standards Secretariat, Acoustical Society of America 1994. ANSI S1.1-1994 (R2004) American National Standard Acoustical Terminology, (12.41) Acoustical Society of America, Melville, NY.
- STEPHEN, F. A. 1996. Movement of velum during speech and singing in classically trained singers. *Journal of Voice* 11/2. 212–221.
- STEVENS, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17. 3–45.
- STEVENS, Kenneth N. 1998. *Acoustic phonetics*. The MIT Press, Cambridge, Massachusetts–London.
- STEVENS, Kenneth N. 2002. Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America* 111/4. 1872–1891.
- STEVENS, Kenneth N. – HOUSE, Arthur S. 1961. An acoustical theory of vowel production and some of its implications. *Journal of Speech, Language and Hearing Research* 4/4. 303–320.
- STRANGE, Winifred – VERBRUGGE, Robert R. – SHANKWEILER, Donald P. – EDMAN, Thomas R. 1976. Consonant environment specifies vowel identity. *Journal of the Acoustical Society of America* 60/1. 213–224.
- SUNDBERG, Johan 1972. An articulatory interpretation of the ‘singing formant’. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 13/1. 45–53.

- SUNDBERG, Johan 1975. Formant technique in a professional female singer. *Acta Acustica united with Acustica* 32/2. 89–96.
- SUNDBERG, Johan 1977. The acoustics of the singing voice. *Scientific American* 236/3. 82–91.
- SUNDBERG, Johan 1979. Perception of singing. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 20/1. 1–48.
- SUNDBERG, Johan 1987. *The science of the singing voice*. Northern Illinois University Press, DeKalb Illinois.
- SUNDBERG, Johan 2001. Level and center frequency of the singer's formant. *Journal of Voice* 15/2. 176–186.
- SUNDBERG, Johan – ASKENFELT, Anders 1981. Larynx height and voice source. A relationship? *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 22/2–3. 23–36.
- SUNDBERG, Johan – LĀ, Filipa – GILL, Brian P. 2011. Professional male singers' formant tuning strategies for the vowel /a/. *Logopedics Phoniatrics Vocology* 36/4. 156–167.
- SUNDBERG, Johan – LĀ, Filipa – GILL, Brian P. 2013. Formant tuning strategies in professional male opera singers. *Journal of Voice* 27/3. 278–288.
- SUNDBERG, Johan – LEANDERSON, R. – VON EULER, Curt 1988. Activity relationship between diaphragm and cricothyroid muscles. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 29/2–3. 83–91.
- SUNDBERG, Johan – NORDSTRÖM, P-E. 1976. Raised and lowered larynx – the effect on vowel formant frequencies. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 17/2–3. 35–39.
- SUNDBERG, Johan – SKOOG, Jörgen 1995. Jaw opening, vowel and pitch. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 36/2–3. 43–50.
- SYRDAL, Ann K. – GOPAL, H. S. 1986. A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America* 79. 1086–1100.
- SZASZÁK György 2010. A beszéd számítógépes feldolgozása. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest, 209–243.
- SZENDE, Tamás 1994. Illustrations of the IPA: Hungarian. *Journal of the International Phonetic Alphabet* 24/2. 91–94.
- TAJIMA, Keiichi – TANAKA, Kuniyoshi – MARTIN, Andrew – MAZUKA, Reiko 2013. Is the vowel length contrast in Japanese exaggerated in infant-directed speech? *Proceedings of Inter-Speech 2013*. 3211–3215.
- THORPE, William C. – CALA, Stephen J. – CHAPMAN, Janice – DAVIS, Pamela J. 2001. Patterns of breath support in projection of the singing voice. *Journal of Voice* 15/1. 86–104.
- TITZE, Ingo R. 1994. *Principles of voice production*. Prentice-Hall, Englewood Cliffs, NJ.
- TITZE, Ingo R. 1998. Voice research: The wide pharynx. *Journal of Singing* 55/1. 27–28.
- TITZE, Ingo R. 2001. Acoustic interpretation of resonant voice. *Journal of Voice* 15/4. 519–528.

- TITZE, Ingo R. 2004. A theoretical study of F_0 - F_1 interaction with application to resonant speaking and singing voice. *Journal of Voice* 18/3. 292–298.
- TITZE, Ingo R. 2008. Nonlinear source-filter coupling in phonation: Theory. *Journal of the Acoustical Society of America* 123/5. 2733–2749.
- TRAUNMÜLLER, Hartmut 1981. Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America* 69/5. 1465–1475.
- VÁRADI Marianna 2010. *A művészi hangképzés kérdései, és néhány példa a magyar operafordítások énekelhetőségének problematikájáról*. Doktori értekezés. Liszt Ferenc Zeneművészeti Egyetem, Budapest.
- VENNARD, William 1964. An experiment to evaluate the importance of nasal resonance in singing. *Folia Phoniatrica* 16. 146–153.
- VENNARD, William 1967. *Singing, the mechanism and the technique*. Carl Fisher, New York.
- VICSÍ Klára 2010a. A hallási folyamat. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest, 27–38.
- VICSÍ Klára 2010b. A beszéd fizikai jellemzése. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest, 38–56.
- VORPERIAN, Hourii K. – KENT, Ray D. 2007. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research* 50. 1510–1545.
- WASSINK, Alicia Beckford – WRIGHT, Richard A. – FRANKLIN, Amber D. 2007. Intraspeaker variability in vowel production: An investigation of motherese, hyperspeech, and Lombard speech in Jamaican speakers. *Journal of Phonetics* 35/3. 363–379.
- WEISS, Rudolf – BROWN, W. S., Jr. – MORRIS, Jack 2001. Singer's formant in sopranos: Fact or fiction? *Journal of Voice* 15/4. 457–468.
- WOLFE, Joe – GARNIER, Maëva – SMITH, John 2009. Vocal tract resonances in speech, singing and playing musical instruments. *Human Frontier and Science Journal* 3. 6–23.
- WOOD, Sidney 1975. The weakness of the tongue-arching model of vowel articulation. *Lund Working Papers* 11. 55–107.
- WRIGHT, James 1975. Effects of vowel nasalization on the perception of vowel height. In FERGUSON, C. A. – HYMAN, L. M. – OHALA, John J. (eds.): *Nasalfest: Papers from a Symposium on Nasals and Nasalization (Language Universals Project)*. Stanford University, Stanford, 373–388.
- YANAGISAWA, Eiji – ESTILL, Jo – MAMBRINO, Lawrence – TALKIN, David 1991. Supraglottic contributions to pitch raising. Videoscopic study with spectroanalysis. *Annals of Otolaryngology, Rhinology, and Laryngology* 100/1. 19–30.
- YANAGISAWA, Eiji – K MUCHA, Steven T. – ESTILL, Jo 1990. Role of the soft palate in laryngeal functions and selected voice qualities. *Annals of Otolaryngology, Rhinology, and Laryngology* 99/1. 18–28.

- ZENKER, Wolfgang – ZENKER, Adolf 1960. Über die Regelung der Simmelippenspannung durch von aussen eingreifende Mechanismen. *Folia Phoniatica* 12/1. 1–36.
- ZWICKER, E. – TERHARDT, E. 1980. Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. *Journal of Acoustical Society of America* 68/5. 1523–1525.

Production and perception of Hungarian vowels in high-pitched soprano singing

The research presented in this volume is based on the apparent contradiction between the requirements of vowel articulation and the requirements of the production of high-pitched singing voice. In accordance with this contradiction, both the production and the perception of sung vowels are examined on Hungarian vowels in the experiments presented in this book.

Due to the well-known relationships between vowel resonances, a.k.a. formants, and the articulatory features of vowels, the first formant (F_1) of close vowels may be relatively low in frequency, or, by all means, lower than that of the fundamental frequency (f_0) or pitch that is often used by soprano singers. Therefore the question arises: what happens if sopranos sing at an f_0 that is higher than the F_1 of the vowel in speech? Does F_1 disappear from the spectrum of the vowel? This scenario is very unlikely, since it would result in a great loss of acoustic energy (i.e., loudness), in a drastic change in timbre, and it would also mean that certain vowel qualities also disappear at high f_0 s. In the opera performances, however, we do not encounter most of these phenomena: the singing voice of a soprano is loud (without any further amplification), homogenous in timbre, and the sung text is also seemingly comprehensible. Hence, in high-pitched singing the modification of vowel articulation is presumed, more specifically, it is assumed that the F_1 is tuned to (or slightly above) the raised high f_0 (this strategy is referred to as $F_1 : f_0$ for short). However, it is very difficult to gather empirical evidence to support this assumption due to the spectral undersampling of the vocal tract transfer function at high f_0 . Moreover, since the $F_1 : f_0$ strategy changes the articulatory configuration of certain vowels at a high f_0 (and the acoustic manifestation of vowels also changes as a result of the spectral undersampling effect), the question arises, how high-pitched sung vowels are perceived. The present volume aims at answering these questions through the description of the acoustic and perceptual properties of high-pitched sung vowels in Hungarian.

Chapter 1 provides an overview of the speech organs and their function in speech, followed by the discussion of the articulatory, acoustic and perceptual features of spoken and sung vowels explored in the literature so far. *Chapter 2* to *Chapter 7* reports on the author's experimental research on the acoustic realization and the perception of sung vowels. Finally, *Chapter 8* presents a general discussion. The main results of the presented research can be summarized as follows:

1. Our results verified that sopranos tend to tune the F_1 to the raised high f_0 , if the f_0 would otherwise exceed the F_1 of the vowels. Accordingly, the lower limit of $F_1 : f_0$ tuning was found to be dependent on the degree of openness or the F_1 of the vowels in speech. Furthermore, it was also found that (as a result of the $F_1 : f_0$ tuning) vowels became more and more open in

production, and the vowel space reduced and shifted towards the position of /a:/ as the f_0 increased. The $F_1 : f_0$ tuning was demonstrated in the entire set of the Hungarian vowels.

2. We successfully measured resonance frequencies in high-pitched sung vowels independently of the voice source by the use of inverse filtering of the acoustic output together with an electroglottogram and by means of external excitation (i.e., the analysis of Electrolarynx recordings).

3. We falsified the previously suggested claim that the opening tendency in the production of sung vowels (accompanying f_0 -increase) is also accompanied by an opening tendency in perception. The only generalizable tendency that was found to be characteristic of the perception of sung vowels with f_0 -increase was a gradually increasing ambiguity of vowel qualities. The tendencies found were most probably not the result of a perceptual opening tendency accompanying the opening tendency in production (as suggested by previous literature), but the result of the decreased tonotopic distance between F_1 and f_0 , and the $F_1 : f_0$ tuning (see below).

4. We only found a notable increase of open vowels /b a:/ as responses when the relative distance between the f_0 and the first harmonic ($2f_0$) and their enhancement were similar to that of the frequency and energy of the first two prominent components of /a:/ in speech. This requires both the frequency of the f_0 to be around 900-1000 Hz and the first two partials ($f_0, 2f_0$) to coincide with the first two resonances (F_1, F_2) which means that the partials are both enhanced by formants. In our studies we demonstrated these correspondences at the musical note f'' (988 Hz).

5. Below 1 kHz the interdependencies of production and perception were not found to be characterized by direct and linear relations. The most dominant tendency found in perception below 5-600 Hz was the identification of the intended vowels as more close vowels; the ratio of more open responses only started to increase above this f_0 . Based on our results we suggested that the percept of more open vowels is not the direct consequence of the opening tendency in production, and it requires both the $F_1 : f_0$ tuning and the spectral undersampling of high-pitched sonorant sounds (resulting also in the correspondence of F_2 and $2f_0$).

6. We suggested that it is also the result of the $F_1 : f_0$ tuning that vowels produced with the second and third degree of openness may appear in perception in high-pitched singing. Without the $F_1 : f_0$ tuning it is only the appearance of vowels with the first and fourth degree of openness that is predicted by the literature of speech perception.

7. We concluded that the consonantal context does not support the identification of the intended vowels in singing through transitional acoustic cues as suggested by previous literature, and the same applied to the voicing onset of isolated vowels (first suggested by the author). Based on the results we concluded that the identification of the intended vowel quality may only be supported by probability: the identification may only be supported by context (vowel embedding in real words and sense texts) and phonology (phonotactics and phonological neighborhood density) which together lead to low or high probability of the carrier words and the intended vowel quality. This claim is also supported by the finding that despite

their experience in production, singers did not exhibit any perceptual advantage in the identification of the intended vowel qualities (see below).

8. We concluded that singers and non-singers identify sung vowels according to very similar perceptual tendencies, and in spite of their advantage in production, singers do not exhibit any advantage in the identification of the intended vowel qualities.

9. We concluded that the modification of the acoustic structure of sung vowels increases with the increasing f_0 , and this cannot be compensated for only by bottom-up perceptual processes; the compensation also requires the activation of top-down perceptual processes (involving syntactic-semantic and pragmatic processing).

A
BESZÉD • KUTATÁS • ALKALMAZÁS
című sorozat eddig megjelent kötetei:

MARKÓ ALEXANDRA:
Az irreguláris zöngé funkciói a magyar beszédben
ISBN 978-963-312-195-5

BÓNA JUDIT:
A spontán beszéd sajátosságai az időskorban
ISBN 978-963-312-199-3

HORVÁTH VIKTÓRIA:
Hezitációs jelenségek a magyar beszédben
ISBN 978-963-312-205-1

NEUBERGER TILDA:
A spontán beszéd sajátosságai gyermekkorban
ISBN 978-963-312-204-4

BEKE ANDRÁS:
Gépi beszélődetektálás magyar nyelvű spontán társalgásokban
ISBN 978-963-312-234-1

DEME ANDREA:
Magánhangzók ejtése és észlelése a szopránéneklésben
ISBN 978-963-312-261-7

A (magas alaphfrekvenciájú) szoprán éneklésben a magánhangzók azonosítása a hallgatók számára nem könnyű feladat. Ennek az az oka, hogy a magánhangzó artikulációjának bizonyos értelemben ellentmond a magas alaphang létrehozására való törekvés az énekes részéről. A magyar köznyelvi magánhangzóknek a szoprán éneklésben jellemző sajátosság megvalósulásának leírására e könyv szerzőjének kutatásai előtt még nem született tudományos igényű vizsgálat. Ugyanakkor ez a könyv az énekelve ejtett magánhangzók észlelésének nemzetközi szakirodalmában is újdonságot hoz.

A szerző alaposan utánajárt a szakirodalomban olvasható megállapítások ellentmondásainak, nagy szorgalommal és alaposítással törekedett az olykor következtetlenül kidolgozott kísérleti módszertanokból előállt féligazságok helyreigazítására. Saját, kontrollált kísérleteit a korábbi kutatások tanulságaiból okulva, ugyanakkor módszertani innovációkat is bevezetve végezte el, a magyar nyelv teljes hangzókészletét figyelembe véve. Az itt bemutatott kutatássorozat mindezek miatt nemcsak világszínvonalú, de a témakörben a létező legkorszerűbb tudást is bemutatja az olvasónak.

ISBN 978-963-312-261-7



9 789633 122617